Using a survey of social and emotional learning and school climate to inform decisionmaking

Tim Kautz, Kathleen Feeney, Hanley Chiang, Sarah Lauffer, Maria Bartlett, and Charles Tilley

December 2021*

*This version adds supplemental analyses to the version released in August 2021. Additions are indicated by a dagger (†).

Using a survey of social and emotional learning and school climate to inform decisionmaking

Tim Kautz, Kathleen Feeney, Hanley Chiang, Sarah Lauffer, Maria Bartlett, and Charles Tilley **December 2021***

The District of Columbia Public Schools (DCPS) has prioritized efforts to support students' social and emotional learning (SEL) competencies, such as perseverance and social awareness. To measure students' SEL competencies and the school experiences that promote SEL competencies (school climate), DCPS began administering annual surveys to students, teachers, and parents in 2017/18. DCPS partnered with the Mid-Atlantic Regional Educational Laboratory to study how the district could use these surveys to improve students' outcomes. The study found the following:

- Students' SEL competencies and school experiences are the most favorable in elementary school and the least favorable in middle school and the beginning of high school. This pattern suggests that schools might provide targeted supports before or during grades 6–10 to promote SEL competencies and school experiences when students need the most support.
- The trajectories of students' SEL competencies and school experiences differed in different schools, to a similar degree as trajectories in academic measures like test scores. To understand why changes in SEL competencies and school experiences differ across schools, DCPS could explore differences in practices between schools with better and worse trajectories. In addition, DCPS could provide targeted support to schools with lower levels of positive change.
- Of the SEL competencies and school experiences in DCPS's survey, self-management—how well students control their emotions, thoughts, and behavior—is most related to students' later academic outcomes. Programs or interventions that target self-management might have the most potential for improving students' outcomes compared to those that target other SEL competencies or school experiences.
- In statistical models designed to predict students' future academic outcomes, SEL competency and school experience data add little accuracy beyond prior academic outcomes (such as achievement test scores and attendance) and demographic characteristics. Prior academic outcomes and demographic characteristics predict later outcomes with a high degree of accuracy, and they may implicitly incorporate the SEL competencies and school experiences. These findings suggest that DCPS would not need to use SEL competencies and school experiences to identify whether or not students are at risk of poor academic outcomes.
- Student, teacher, and parent reports on SEL competencies and school experiences are positively related across schools, but they also exhibit systematic differences, suggesting that some respondent groups may not be aligned in their view of SEL competencies and school experiences. These differences may serve as a tool to help DCPS target efforts to improve communication among students, teachers, and parents.

Why this study?

Students' social and emotional learning (SEL) competencies—such as how well students persevere, manage their thoughts and emotions, and understand what others think and feel—have been shown to be related to many life outcomes and, importantly, can be shaped through education. Such SEL competencies rival cognitive measures (such as IQ) in predicting long-term outcomes, including educational attainment, health, earnings, and employment (Heckman & Kautz, 2012). They are also related to shorter-term outcomes, such as academic achievement (O'Conner et al., 2017). In addition, interventions can improve SEL competencies throughout grades

K–12, suggesting that SEL competencies are a promising avenue through which education can improve the long-term success of students (Elango et al., 2016; Kautz et al., 2014). Evidence suggests that one way schools may be able to boost SEL competencies is by fostering a positive school climate—the tangible and intangible attributes of a school that support students' development—including relationships among students and staff, school discipline, student engagement, and safety (SRI International, 2018).

Inspired by this type of evidence, the District of Columbia Public Schools (DCPS) has prioritized supporting positive SEL outcomes (for example, perseverance) and has begun administering surveys to measure students' SEL competencies and school climate. In its 2017–2022 Strategic Plan—a proposal that DCPS uses to outline key goals and hold itself accountable to the public—DCPS highlighted SEL competencies and school climate as key components in its priority to "educate the whole child" (DCPS, 2017). All six goals reflect this priority, either explicitly or implicitly. For example, DCPS has set an explicit goal that, by 2022, 100 percent of students will feel "loved, challenged, and prepared," as measured by an index based on SEL competencies and school climate (DCPS, 2017; see figure A1 in appendix A). Other goals—such as improving college and career readiness and increasing re-enrollment rates—are implicitly related to SEL competencies and school climate in that they might be furthered by better identifying at-risk students and boosting their SEL competencies and school experiences. To track progress toward its goals, DCPS launched annual surveys (developed by Panorama Education) in spring 2018 that collect information from students, teachers, and parents on SEL competencies and experiences that reflect school climate.

Despite a strong research base supporting the importance of SEL competencies, education stakeholders require additional guidance on how to use this kind of survey data to inform and improve their programming and interventions. Because of this need, as well as their ambitious goals and priorities, DCPS partnered with Regional Educational Laboratory (REL) Mid-Atlantic in examining how their existing data might be used to inform education decisions and improve key student outcomes, including those highlighted in their 2022 goals. By exploring avenues to identify students who could benefit from additional support—such as those who report not feeling loved, challenged, and prepared—the findings may suggest how DCPS can best target its growing investments in SEL and school climate. Because improvement practices differ across schools, a better understanding of its SEL and school climate data would allow the district to enhance the design, consistency, and targeting of its practices and programming to attain its goals. Furthermore, DCPS might better identify and serve students at risk of poor outcomes and improve the quality and ratings of low-performing schools.

Box 1. Key terms

Academic measures. The study used academic measures that serve as both predictors and student outcomes, which include academic proficiencies (such as proficiency/college readiness in math and English language arts) and academic behaviors (such as attendance, suspensions, grade progression, and credits earned toward graduation) (see table B3 in appendix B).

Classification accuracy. The percentage of students whose outcomes are correctly classified by a statistical model.

Demographic variables. The study used demographic data, which include students' gender, special education status, English learner status, grade level, and race/ethnicity, as well as the District of Columbia Public Schools' (DCPS) classification of whether the student is at risk (see table B3 in appendix B).

Predictive power. The strength of association between a predictor or group of predictors and student outcomes.

School climate scales. DCPS uses a customized version of Panorama Education's survey of school climate and social and emotional learning (SEL) competencies. The school climate scales include items that ask students, parents, and teachers about their personal experiences at the school (see table B1 in appendix B). When referring to individual respondents' reports, this study uses the term school experiences. Averaging the measures of school experiences across respondents at a school provides a measure of a school's climate. Response options are Likert scales relating to the question text (such as agreement or frequency), with answers of (1) indicating a low level and (5) indicating a high level. The survey covers three components of school experiences:

- *Rigorous expectations*: A 1- to 5-point scale based on items designed to capture "how much students feel that their teachers hold them to high expectations around effort, understanding, persistence, and performance in class." [a]
- *Sense of belonging*: A 1- to 5-point scale based on items designed to capture "how much students feel that they are valued members of the school community." [a]
- *Student satisfaction*: A 1- to 5-point scale based on items developed by DCPS to capture student reports of how satisfied they are with their school experience. [b]

SEL competency scales. The survey also covers four SEL competencies:

- *Perseverance* (also called *grit*)*:* A 1- to 5-point scale based on items designed to capture "how well students are able to persevere through setbacks to achieve important long-term goals (not limited to academics), taking into account their experiences and identities." [a]
- *Self-management:* A 1- to 5-point point scale based on items for all students designed to capture "how well students manage their emotions, thoughts, and behaviors in different situations." [a]
- *Self-efficacy:* A 1- to 5-point scale based on items designed to capture "how much students believe they can succeed in achieving academic outcomes." [a]
- *Social awareness:* A 1- to 5-point scale based on items designed to capture "how well students consider the perspectives of others and empathize with them." [a]

Student Loved, Challenged, and Prepared Index. The study used a student outcome developed to measure progress toward DCPS's strategic goal of having 100 percent of students feeling loved, challenged, and prepared. The Index indicates whether or not students feel loved, challenged, and prepared and is based on student reports of perseverance, self-management, self-efficacy, sense of belonging, and rigorous expectations (see table B3 in appendix B; see background on the development of the Index in appendix A).

Note

a. Panorama Education, n.d.

b. The student satisfaction items have four response categories. When reporting findings, the study team rescaled the student satisfaction scale to range from 1 to 5, so it was comparable to the other scales. See appendix B for details on the rescaling.

Research questions

This study addresses four key research questions that will help DCPS understand and use measures of SEL competencies and school experiences.

Research question 1. How do average SEL competencies and school experiences differ across grade levels and change for individual students between years? Do student and teacher reports of SEL competencies and school experiences change in similar ways across grade levels? To what extent do the average differences in students' SEL competencies and school experiences across grades differ by the type of students (such as students classified by gender, race/ethnicity, and academic achievement)? How are individual students' reports of SEL competencies and school experiences associated between years, and how does that association compare to that of other variables (such as achievement test scores, absences, and suspensions)?

To identify and support students who need the most help, DCPS requires a greater understanding of whether students enrolled in some grades tend to have lower SEL competencies and have poorer school experiences and how much students' reports of SEL competencies and school experiences change between years. Evidence from other districts suggests that students' self-reported SEL competencies do not necessarily improve—and can even decline—across grades (West et al., 2018). If DCPS's data reveal a similar decline, then DCPS might focus its SEL efforts on grades before or during the times when students tend to struggle most. Similarly, if students in some grades have worse school experiences, DCPS might focus school climate improvement efforts on those grades. Finally, if some subgroups of students struggle more than others, DCPS could target supports to them. Information

on how individual students' SEL competencies and school experiences relate between successive years provides a sense of the stability of the measures over time.

Research question 2. To what extent do year-to-year changes in individual students' SEL competencies and school experiences differ across schools?

The extent to which year-to-year changes in students' SEL competencies and school experiences differ across schools can provide a basis for exploring differences in practices across schools and shed light on the promise of targeting supports to specific schools. If year-to-year changes vary across schools, then students in some schools have, on average, relatively more positive changes in SEL competencies and school experiences than students in other schools. In this case, DCPS may wish to explore the SEL- and climate-related practices of all schools, observing whether some practices are consistently associated with more positive outcomes than others. At the same time, other factors could matter as well. For example, if family or community supports tend to be greater at some schools than others, then school practices might not account for the systematic differences in year-to-year changes in SEL competencies and school experiences across schools. Regardless of the source of any differences, schools with little or no positive change might benefit from additional supports to help improve students' growth in SEL competencies and school experiences.

Research question 3. How do measures of SEL competencies and school experiences relate to future outcomes, and how do they complement other available data for predicting future outcomes? To what extent do individual SEL competencies and school experiences relate to student outcomes measured one and two years later (such as achievement test scores, absences, suspensions, and whether a student feels loved, challenged, and prepared)?[1] When other data are available—such as demographic information, achievement test scores, absences, and suspensions—to predict students' future outcomes, to what extent does adding measures of SEL competencies and school experiences improve the predictive power and accuracy of those predictions? Which types of data and statistical models could best help DCPS classify whether students are at risk of having negative outcomes?

Information about the relationships between current SEL competencies and school experiences and future outcomes can help DCPS prioritize which SEL competencies and aspects of school climate to focus on, as well as refine ways to identify students at risk of poor future outcomes. For many interventions and initiatives, the primary goal is to improve students' SEL competencies and school experiences in ways that also boost meaningful longer-term outcomes, such as grade progression and high school graduation. Similarly, for identifying struggling students, a primary goal is to use proximate data to identify students at risk of performing poorly on later outcomes, like dropping out of school. Research question 3 explores the predictive power of the SEL competencies and school experiences and compares their predictive power to that of other administrative data.

These analyses will inform DCPS in two distinct ways. First, they will suggest which SEL competencies and student experiences DCPS might target to improve longer-term student outcomes. Given that DCPS is already investing in this area, the results will inform which of the competencies and school experiences to prioritize. Second, they will suggest how DCPS can best identify students who might be at risk of falling behind in the future and whether data on SEL competencies and school experiences can improve these predictions. These issues are separate, because it is possible that measures of SEL competencies and school experiences relate to future outcomes but they would not improve the ability to predict outcomes beyond using other available data. The comparison of different statistical models will also provide practical guidance to DCPS on the benefits of different approaches to prediction.

[1] Supplementary analyses explore the extent to which year-to-year changes in SEL competencies and school experiences relate to each other and changes in student outcomes (see appendix D for details).†

Subgroup analyses of grades 3 and 8 will inform specific teams at DCPS.[2] Analyses of grade 3 students will suggest to DCPS's early literacy team which SEL competencies and school experiences in elementary school are most critical for early English language arts (ELA) achievement. Analyses of grade 8 students will help DCPS's graduation team to understand which SEL competencies and school experiences in middle school enable students to successfully transition to high school, a particularly challenging transition for many (Benner, 2011).

Research question 4. How do measures of perseverance and rigorous expectations align across students, parents, and teachers? Across schools, to what extent do survey reports on these measures from students, parents, and teachers align? Is alignment associated with characteristics of schools (such as the demographic characteristics of their student population and the schools' accountability ratings) and response rates on the survey?

Information on the extent to which respondents' reports align on these two measures will allow DCPS to better understand how to address shortfalls in students' SEL competencies and school experiences and improve communication among students, parents, and teachers. After reviewing the 2017/18 survey data, school staff struggled to understand why some students reported low levels of SEL competencies and school experiences. If reports across respondents are misaligned, it might help the district pinpoint the cause of these gaps, particularly around communication and engagement with available programming. For example, if teachers at a school indicate that students are prepared but the students report that they do not feel prepared, then school leaders might take a different action than if teachers and students agreed that the students are not prepared. Research question 4 addresses this by examining alignment between school-level measures of SEL competencies and school experiences across the three respondent types. The results may also help DCPS interpret results from school-level tracking and determine whether including information from different types of respondents provides different information.

The data sources, sample, and methods used to answer these questions are described in box 2 and appendix B.

Box 2. Data sources, sample, and methods

Data sources. The key data sources for this study are DCPS's Panorama Education student, teacher, and parent surveys and administrative records from the 2017/18, 2018/19, and 2019/20 school years (see, for example, DCPS, 2019). The COVID-19 pandemic affected the collection of the survey data from the 2019/20 school year (see appendix B), which limited its use in this study. Our main analyses, therefore, did not include the 2019/20 survey data. The administrative data covered a range of student demographic characteristics and academic measures, as well as school characteristics (see appendix B).

Sample. During the 2017/18–2018/19 school years, there were 39,791 unique DCPS students enrolled in grades 3 through 12 (see table B2 in appendix B). Of these students, 30,462 responded to the Panorama Education survey in one or more years and could therefore be included in the analyses. The study sample included 4,273 unique teacher respondents and 12,216 parent responses on the SEL surveys during the 2017/18–2018/19 school years (no 2019/20 data were used in these analyses). Sample sizes and number of observations for each analysis are in table B4.

Methodology.

Nonresponse analyses. The study team assessed survey nonresponse bias in the student and teacher surveys (see appendix B). The results suggested the potential for nonresponse bias, which could lead to findings that do not represent the student and teacher populations across DCPS. As a result, the analyses of student and teacher data used nonresponse weights based on administrative data that were available for students and teachers regardless of whether they completed the survey. Due to a lack of data on parents who did not respond, the study team did not assess nonresponse bias on the parent survey.

Research question 1. To inform how measures of SEL competencies and school experiences differ across grades, the study team calculated the average level of each measure reported by students and teachers by grade levels within the two academic

[2] To provide additional context, the study team also explored the link between credits behind in grade 9 and high school graduation, discussed in appendices B and C.

years included in the sample. The student reports were averaged within individual grades. Because a teacher's report could apply to students in multiple grades, the teacher reports were averaged within ranges of grades based on the level of their school: grades K–5 (elementary), grades 6–8 (middle), or grades 9–12 (high). To assess whether the measures varied across grades and school levels, the study team conducted *F*-tests of the null hypothesis that the measures were equal across grades and school levels. These analyses compared students in different grades when the survey was administered. To understand how individual students developed between years, the study team estimated the correlation between student reports of the same measure over time (the year-to-year correlation) and between other variables like test scores, absences, suspensions.

Research question 2. The study team examined the extent to which changes in students' SEL competencies and school experiences differed across schools. For each measure, the study team calculated the expected difference in percentile points between students who attend a school with a high level of (positive) change compared to a school with an average level of change. A high level of change is defined as one standard deviation above average within DCPS. As described in appendix B, the calculation was based on the intraclass correlation coefficient (ICC) of year-to-year changes in students' measures. The ICC represents the fraction of a measure's total variance that arises across schools compared to within schools. A higher percentile point difference (and ICC) indicates that changes in students' measures vary more across schools. As a benchmark, the analyses also examined changes in achievement test scores, suspensions, and in-seat attendance, as well as the variation across schools for measures during a fixed year (as opposed to changes across years).

Research question 3. To describe how individual measures of SEL competencies and school experiences relate to later outcomes, the study team estimated pairwise correlations between each measure and outcome. If the pairwise correlation is positive, then, on average, when one variable takes a higher value, the other variable takes a higher value. Similarly, multivariate correlations (the square root of the adjusted *R*-squared statistic) provided evidence on the extent to which groups of measures predicted student outcomes. The higher the multivariate correlation, the more related the variable is to the group of other variables. This correlation also suggests how predictions would improve when variables are added to a group of predictors. If adding a variable to a group of predictors increases the multivariate correlation, then the variable provides additional predictive power above and beyond the initial group. The groups included students' SEL competencies and school experiences, demographic variables, academic measures (such as in-seat attendance), and all three types of predictors combined. By comparing the results from these groups of predictors, the study team assessed the extent to which the measures of SEL competencies and school experiences added predictive power above and beyond the academic and demographic variables that DCPS previously collected.

The study team conducted additional analyses to provide information on how well DCPS could use different data sources to identify whether or not individual students are at risk of poor outcomes. The study team recoded the continuous outcomes used in the correlational analyses to be dichotomous outcomes (for example, whether or not a student was chronically absent). Probit and machine learning models (random forests) were used to calculate how frequently different groups of predictors classified students' future outcomes accurately. These analyses complement the correlational analyses by placing the results in terms of the accuracy of identifying students who are at risk of having negative future outcomes.

Research question 4. Finally, the study team conducted two analyses to examine the alignment among students, teachers, and parents in their reports of perseverance (an SEL competency) and rigorous expectations (a school experience)—the two measures available in all three surveys. Both analyses involved examining school-level averages of the reports from the three types of respondents, as teacher and parent responses cannot be linked to individual students. First, the study team estimated pairwise correlations between the school-level averages among the respondent types. Second, the study team compared the averages between school-level reports across respondent types.

Due to a lack of data on parents who did not respond to the survey, these analyses cannot account for nonresponse in parent reports (see appendix B). For consistency, the study team did not use nonresponse weights when calculating the school-level averages of student and teacher results. The results, therefore, generalize to the population of students, teachers, and parents who responded to the survey, as opposed to the full population.[3]

[3] Supplemental analyses reveal that adjusting for nonresponse in the student and teacher reports made little difference, suggesting that the results for these two groups may generalize to the full population. However, given the low estimated response rate among parents—approximately 12 and 15 percent in 2017/18 and 2018/19, respectively (DCPS, 2019)—the study team strongly cautions against generalizing the analyses that include parent reports.

Benchmarks. This study uses two benchmarks when describing the strength of correlations and differences in SEL competencies and school experiences between groups of students.

1. **Correlations.** Based on past evidence on the extent to which cognitive (IQ and achievement) tests predict other academic outcomes, this study describes ranges of correlations as follows: 0.0 to 0.09 is low; 0.10 to 0.19 is moderate; 0.20 to 0.29 is substantive; and 0.30 and above is high (see appendix B for details). These descriptions also apply when describing the difference between two correlations.
2. **Comparing SEL competencies and school experiences between students.** Based on past evidence on the extent to which school-based SEL programs improve students' SEL competencies and school experiences, this study describes differences in SEL competencies and school experiences between groups of students as follows: 0.0 to 0.09 standard deviations is small; 0.10 to 0.19 standard deviations is moderate; 0.20 to 0.29 standard deviations is substantive; and 0.30 standard deviations and above is large (see appendix B for details).

More information about the study data sources, sample, and methods is in appendix B.

Findings

This section presents findings to address the study's research questions, with additional findings in appendix C.

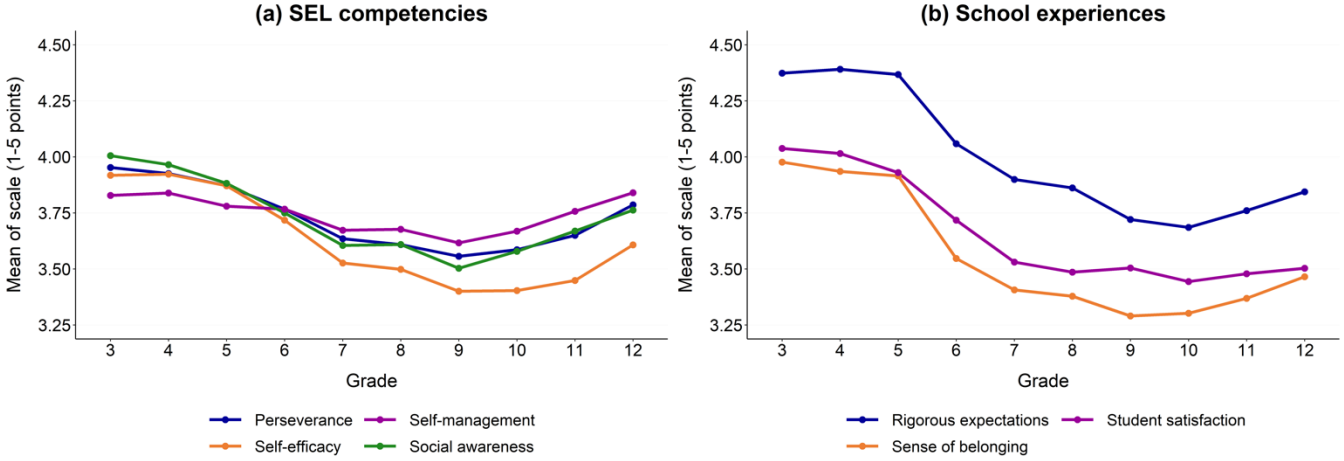
Both student and teacher reports of SEL competencies and school experiences were highest in elementary school and lowest in middle school and high school

Student reports of SEL competencies and school experiences exhibited a U-shaped pattern, with the lowest reports in middle school and the beginning of high school

For all student-reported SEL competencies and school experiences, the average value was highest in elementary school (grades 3–5) and dipped in middle school (grades 6–8) and the first half of high school (grades 9 and 10), at which point the average value began to rise again (figure 1).[4] In addition, formal tests indicated that each measure varied across grades in a statistically significant way. Of all of the measures, rigorous expectations differed the most across grades, with a range of 0.71 on the 5-point scale—representing a large difference in terms of the study's benchmarks.[5] Of the SEL competencies, self-efficacy showed the biggest differences across grades. The general patterns are also consistent with those from the CORE Districts, a network of large urban districts in California (West et al., 2020).[6] Students' average SEL competencies and school experiences also exhibited similar patterns within student subgroups as they do in the full sample (see tables C1–C7 in appendix C). Notably, in nearly every grade level, female students reported moderately higher levels of self-management and social awareness than their male peers (differences that were statistically significant).

[4] The year-to-year correlations between individual SEL competencies and school experiences tended to be lower than those of academic measures, consistent with the findings that the SEL competencies and school experiences evolve substantially across grades (see figure C2).
[5] The patterns were similar when estimating averages by age (see figure C1 in appendix C) and for subgroups of students (see tables C1–C7).
[6] Importantly, these estimates should not be interpreted as how measures of individual students' SEL competencies and school experiences change over time, because they involve comparing different students across cohorts and grades. The grade-level patterns in measures of SEL competencies and school experiences were nearly identical between the two survey years in the study despite coming from different cohorts, suggesting that differences across cohorts played a minimal role (see tables C1–C7 in appendix C). However, the study cannot rule out that the U-shaped pattern emerged due to changes in the composition of enrolled students over time.

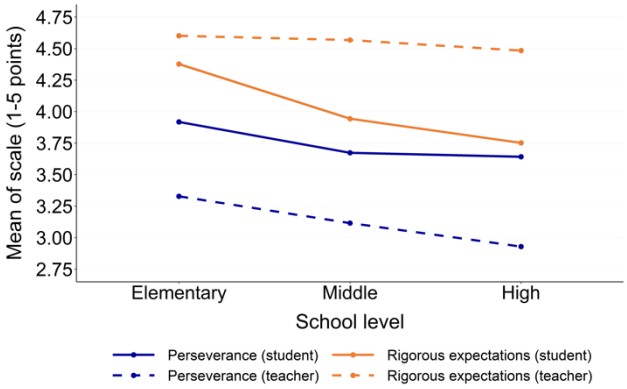
Figure 1. Student self-reports of SEL competencies and school experiences peaked in elementary school, declined in middle school and early high school, and increased at the end of high school



SEL is social and emotional learning.
Note: The figure shows for each grade the mean of each SEL competency and school climate scale (described in appendix B) for the 2017/18 and 2018/19 school years combined. The means were calculated using nonresponse weights, as described in appendix B. *F*-tests of the null hypothesis that the scales were equal across grades are significant at $p < .001$ for each scale.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

For the teacher-reported measures of perseverance and rigorous expectations, the average values were highest in elementary school (grades K–5) and declined through middle school (grades 6–8) and high school (grades 9–12) (represented by the dotted lines in figure 2). In addition, formal tests indicated that each measure varied across school levels in a statistically significant way. Although the teacher reports did not exhibit a U-shaped pattern, they were consistent with the student reports. Because the teacher data were averaged within each school level—that is, elementary, middle, and high school grades—declines within individual grades might be offset by increases in other grades. Supporting this possibility, averaging the student reports to the school level (rather than grade level) produced a similar pattern as the teacher reports, because the student reports of these two measures were lowest in grades 9 and 10 but higher in grades 11 and 12 (represented by the solid lines in figure 2). The steepness of the decline differed between the teacher and student reports, with the decline in perseverance being steepest for teacher reports and the decline in rigorous expectations being steepest for student reports.

Figure 2. School-level teacher and student reports of students' SEL competencies and school experiences declined from elementary school to high school

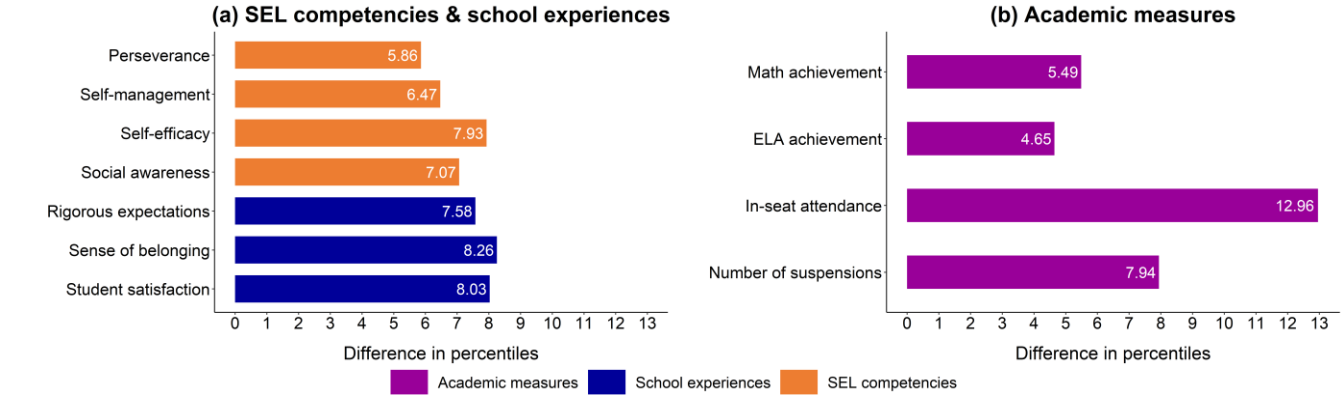


SEL is social and emotional learning.
Note: The figure shows, for each level of school, the mean of each SEL competency and school climate scale (described in appendix B) for the 2017/18 and 2018/19 school years combined. The means were calculated using nonresponse weights, as described in appendix B. *F*-tests of the null hypothesis that the scales were equal across school levels are significant at $p < .001$ for each scale.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Year-to-year changes in students' SEL competencies and school experiences differed across schools, and to a similar degree as year-to-year changes in academic measures

Among schools with high levels of positive change (one standard deviation above average), the average year-to-year change in students' SEL competencies and school experiences tended to be 5.9 to 8.3 percentile points higher than among schools with average year-to-year change (figure 3). This was less than those for changes in in-seat attendance, similar to those for suspensions, and somewhat higher than those for math and ELA achievement test scores.[7, 8] These findings indicate that changes in SEL competencies and school experiences differed across schools to a similar degree as changes in academic measures.[9, 10]

Figure 3. Differences across schools in year-to-year changes in SEL competencies and school experiences were similar to those for academic measures



ELA is English language arts. SEL is social and emotional learning.
Note: The figure shows the difference in percentiles of each measure between schools with average year-to-year improvements and schools with high year-to-year improvements (one standard deviation above average). The measures are defined in table B3 in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

SEL competencies and school experiences were related to students' future outcomes, but their predictive power relative to academic measures varied

Of the SEL competencies and school experiences, self-management was most strongly associated with students' future outcomes

Self-management was the SEL competency or school experience most strongly correlated with six of the nine outcomes included in the analysis, with correlations ranging from 0.09 to 0.23 (which the study considers a low-to-substantive range; see panel a of table 1).[11, 12, 13] For the grade progression and in-seat attendance outcomes, the correlations were strongest (at the moderate range) for rigorous expectations, but other SEL competencies and school experiences were similarly correlated with these outcomes. Finally, whether students felt loved, challenged, and prepared was the most related to sense of belonging—this was likely because whether students

[7] The patterns are similar when considering the ICCs, rather than percentile measures (see figure C3 in appendix C).

[8] The estimates of the ICCs for year-to-year changes in math and ELA achievement test scores are comparable to those compiled in Schochet & Chiang (2010).

[9] Compared to the year-to-year changes in measures, the measures during a fixed year tended to vary more across schools (see figure C4 in appendix C), consistent with existing literature that has examined achievement test scores (Schochet & Chiang, 2010).

[10] The difference in year-to-year changes between average- and high-change schools was substantial relative to the year-to-year change in average schools (see table C9 in appendix C).

[11] The outcomes were all recoded so that a higher value of the outcome is beneficial. For example, a positive correlation between a predictor and the number of suspensions indicates that higher values of the predictor are associated with *fewer* suspensions.

[12] The results were similar when examining the predictive power of outcomes measured one year later (see table C13 in appendix C).

[13] Year-to-year changes in individual SEL competencies and school experiences are highly correlated with each other, but have low correlations with changes in academic measures (see tables D1-D2 in appendix D).†

felt loved is based on items from the sense of belonging scale.[14] With a few exceptions, the remaining measures of SEL competencies and school experiences were also positively related to students' future outcomes at lower levels.

Findings for key subgroups of students of interest to DCPS were consistent with findings from the main sample:

- **Early literacy.** The findings were similar when considering relationships between grade-3 SEL competencies and school experiences and grade-3 ELA achievement, with the strength of the correlations ranging from low to substantive (see table C10 in appendix C). As in the main sample, self-management was the most correlated with grade-3 ELA achievement.
- **Grade 8 predictors of high school success.** Students' SEL competencies and school experiences in grade 8 were moderately predictive of the number of credits they were behind at the end of grade 9 (see table C11 in appendix C). These grade-8 SEL competencies and school experiences are likely related to students' eventual graduation, because the extent to which students are behind in credits in grade 9, in turn, accurately classifies whether students will graduate 75 percent of the time (see table C12).

Of the individual predictors, the best predictor of a later student outcome tended to be an earlier measure of that outcome

Compared to the SEL competencies and school experiences, early academic measures tended to have stronger relationships with future academic outcomes (panel b of table 1). For example, of all predictors, ELA achievement in 2017/18 was the most correlated with ELA achievement in 2019/20, with a high correlation of 0.78. Similarly, for the one outcome based on SEL competencies and school experiences—whether students felt loved, challenged, and prepared—the best individual predictors were SEL competencies and school experiences. Grade progression, graduation, Advanced Placement (AP) credits earned, and credits behind did not have earlier measures based on those outcomes. Of those, in-seat attendance was most strongly related to future grade progression, credits behind, and graduation. ELA achievement was the most strongly related to AP credits earned. For all academic outcomes, the strongest individual correlation was with an academic measure, rather than a measure of an SEL competency or school experience (compare panels a and b in table 1).

As a group, SEL competencies and school experiences did not help predict future academic outcomes more accurately when demographic and academic measures were available, but they did help predict whether students felt loved, challenged, and prepared

As a group, demographic and academic measures strongly predicted future academic outcomes, and adding information on SEL competencies and school experiences did not improve those predictions. Compared to the group of SEL competencies and school experiences, the group of demographics and academic measures better predicted academic outcomes, with high multivariate correlations ranging from 0.33 to 0.87 (see panel d in table 1). For academic outcomes, when using all predictors, the correlation was, at most, 0.01 higher than when using only the demographic and academic measures (compare the last two rows in panel d). Although the SEL competencies and school experiences were moderately to highly predictive of academic outcomes on their own, they added little predictive power because the information they capture was also captured by the demographic and academic predictors. In contrast, adding the SEL competencies and school experiences improved the prediction of whether students felt loved, challenged, and prepared (compare the last two rows in panel d).

[14] In supplementary analyses, the study team examined the correlations between school-level teacher and parent reports of perseverance and rigorous expectations and school-level outcomes one year later (see table C14 in appendix C). Both perseverance and rigorous expectations were correlated with outcomes, but the relative strength of the correlations differed by respondent type. The teacher reports of perseverance were more positively correlated with outcomes than were the reports of rigorous expectations. However, the reverse was true of the parent reports for academic behaviors (suspensions and attendance) and whether students felt loved, challenged, and prepared.

Table 1. Correlations between predictors and students' outcomes measured two years out

Predictor(s) in 2017/18	Outcomes measured two years out in 2019/20: Progressed successfully (grades 3–12)	Number of suspensions (grades 3–12)	Math achievement (grades 9–12)	ELA achievement (grades 9–12)	In-seat attendance (grades 3–12)	Number of AP credits earned (grades 10–12)	Number of credits behind (grades 3–12)	Graduated within 2 years (grade 11)	Loved, challenged, and prepared (grades 3–12)
a. SEL competencies and school experiences									
Perseverance	0.06	0.01	-0.05	-0.05	0.09	-0.04	0.03	0.07	0.22
Self-management	0.09	**0.11**	**0.23**	**0.23**	0.14	**0.18**	**0.12**	**0.16**	0.21
Self-efficacy	0.10	0.01	0.08	0.09	0.14	0.07	0.05	0.06	0.25
Social awareness	0.10	0.07	0.13	0.13	0.17	0.14	0.06	0.09	0.25
Rigorous expectations	**0.12**	0.02	0.04	0.04	**0.17**	0.02	0.06	0.10	0.21
Sense of belonging	0.08	0.01	0.00	0.00	0.13	0.03	0.00	0.04	**0.29**
Student satisfaction	0.09	0.06	0.11	0.11	0.17	0.10	0.05	-0.02	0.24
b. Academic measures									
Math achievement	0.21	0.15	0.77	0.78	0.28	0.46	0.24	0.22	**0.04**
ELA achievement	0.23	0.15	**0.79**	**0.78**	0.30	**0.56**	0.28	0.24	0.01
In-seat attendance	**0.54**	0.08	0.27	0.28	**0.65**	0.28	**0.50**	**0.53**	-0.01
Number of suspensions	-0.20	**-0.29**	-0.21	-0.20	-0.31	-0.18	-0.23	-0.19	-0.03
c. Groups of variables									
SEL competencies and school experiences	0.14	0.14	0.31	0.31	0.21	0.25	0.13	0.16	**0.32**
Demographics	0.45	0.20	0.68	0.67	0.52	**0.61**	0.24	0.21	0.23
Academic measures	**0.54**	**0.31**	**0.83**	**0.83**	**0.67**	0.57	**0.52**	**0.54**	0.05
d. Combinations of groups of predictors									
Demographics and academic measures	0.59	0.33	0.87	0.86	0.71	0.69	0.54	0.55	0.24
All predictors	**0.59**	**0.34**	**0.87**	**0.86**	**0.71**	**0.69**	**0.54**	**0.56**	**0.38**

Absolute value of the correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

AP is Advanced Placement. ELA is English language arts. SEL is social and emotional learning.

Note: The table shows pairwise correlations and multivariate correlations between predictors in the left column and the outcomes in the top row. The outcomes were all recoded so that a higher value of the outcome is beneficial. The bold font indicates the correlation with the highest absolute value within the column and panel. The correlations were calculated using nonresponse weights, as described in appendix B. The sample included students who completed the SEL survey in 2017/18. The math and ELA achievement outcomes were based on high school students' Preliminary SAT (PSAT) and SAT scores.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

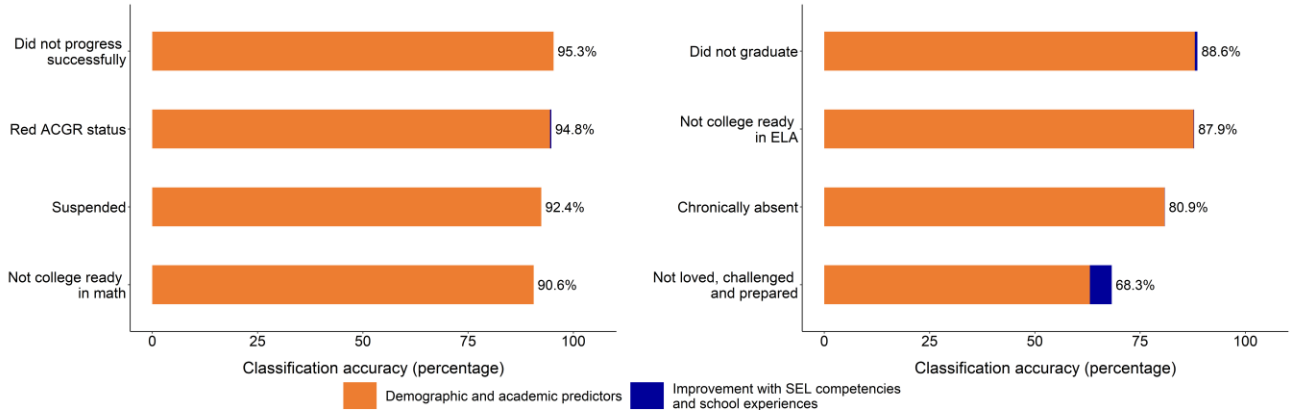
Overall, SEL competencies and school experiences added little value in classifying whether students were at risk of poor academic outcomes beyond demographic and academic predictors

Across all outcomes, the predictive models that used demographic characteristics and academic measures accurately classified students' outcomes between 63.1 and 95.3 percent of the time (figure 4).[15] Consistent with the multivariate correlations, the classification accuracy for academic outcomes improved by at most 0.6 percentage points when adding the SEL competencies and school experience variables to models that included demographic characteristics and academic measures (see table C15 in appendix C). However, adding the SEL competencies and school experiences increased the accuracy by up to 5.2 percentage points when classifying

[15] Complex machine learning algorithms did not systematically perform better compared to probit models when assumptions and modeling approaches were aligned (see table C15 in appendix C).

whether students felt loved, challenged, and prepared two years later—a measure based on future SEL competencies and school experiences. Relative to the accuracy of classifying individual students' outcomes using the typical outcome observed among their peers (that is, the null model), the predictive models improved the classification accuracy for some outcomes more than others (see table C15).[16] For example, the predictive model improved the classification accuracy of college readiness in ELA two years later by 27.8 percentage points. On the other hand, the predictive model improved the classification accuracy of whether students would be suspended two years later by only 0.1 percentage points. Notably, the overall classification accuracy does not distinguish between students who have negative versus positive outcomes. If DCPS places more importance on classifying students who eventually have negative outcomes, DCPS could refine the models to do so (box 3). Additional evidence suggests that these general findings hold for refinements that place more weight on identifying students who eventually have negative outcomes (see figure C5).

Figure 4. Adding SEL competency and school experience predictors improved classification accuracy by at most 5.2 percentage points relative to models that included demographic and academic predictors



ACGR is Adjusted Cohort Graduation Rate. AP is Advanced Placement. ELA is English language arts.
Note: The figure shows the accuracy of predictive models that used demographic variables and academic measures to predict outcomes two years later, as well as the improvement in accuracy by adding SEL competencies and school experiences (described in appendix B). The estimates were based on probit models and calculated using nonresponse weights, as described in appendix B. The measures are defined in table B3.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Box 3. Interpreting classification accuracy

This study estimated how well predictive models accurately classified the future outcomes of individual students. For each outcome, the model generated the probability that each student will have a negative outcome based on a set of predictors. Each student with a probability above a threshold was classified as at risk of having a negative outcome (for example, not progressing to the next grade), and those below the threshold were classified as not at risk of a negative outcome. The study selected the thresholds to maximize the overall classification accuracy—the percentage of individual students who were classified correctly, regardless of whether they eventually had positive or negative outcomes. The overall classification accuracy can be split into two components:

1. **The percentage of students with a negative outcome whom the model accurately classified.** This component measures how accurately the model and threshold classified individual students who eventually had a negative outcome (for example, did not graduate).
2. **The percentage of students with a positive outcome whom the model accurately classified.** This component measures how accurately the model and threshold classified individual students who eventually had a positive outcome (for example, did graduate).

[16] The baseline accuracy is the percentage of students who would be correctly classified if all students were classified with the most prevalent value for each outcome. For example, because 34.9 percent of students were chronically absent two years out, the baseline accuracy was 65.1 percent—the classification accuracy from assuming that all students were not chronically absent. For additional details, see appendix B.

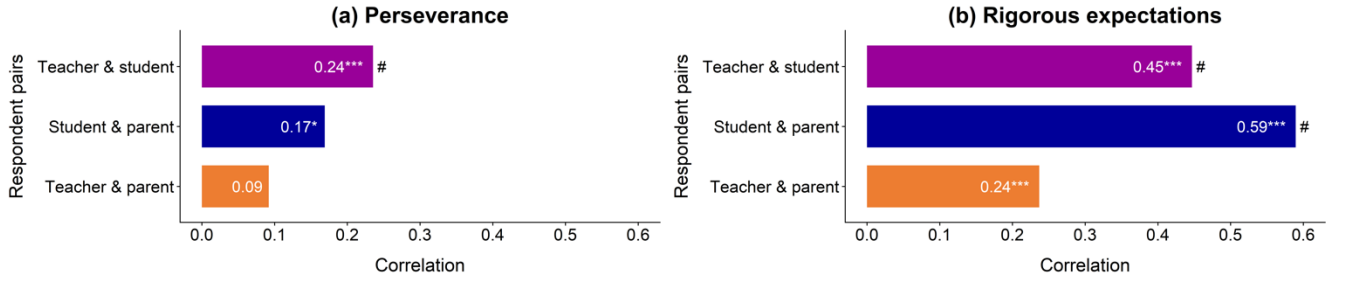
The probability threshold governs a tradeoff between accurately classifying students who eventually have negative and positive outcomes. Reducing the threshold means that students who have a positive outcome are accurately classified a higher percentage of the time, but students with a negative outcome are accurately classified a lower percentage of the time. For some outcomes, it might be more important to accurately classify students who eventually have negative outcomes to ensure resources are allocated to students most likely to require supports, even at the expense of misclassifying some students who eventually have positive outcomes. Refining the threshold can give more weight to such students. A receiver operating characteristic curve illustrates this tradeoff for all possible thresholds and can be used as a tool for determining such refinements (see figure C5 in appendix C).

At the school level, student, parent, and teacher reports of perseverance and rigorous expectations were positively related; however, teacher and parent reports exhibited the greatest differences

When comparing student, parent, and teacher reports of perseverance and rigorous expectations, schools with more favorable reports from one group tended to, but did not always, have more favorable reports from the other groups.[17] The correlation between school-level averages and respondent reports of perseverance and rigorous expectations ranged from 0.09 (low) to 0.59 (high; figure 5). This range is comparable to estimates from other studies that have explored the alignment of student, parent, and teacher reports of individual students' SEL competencies (Barbaranelli et al., 2003; Elliott et al., 2020). The correlations among respondent types for rigorous expectations were substantially higher than those for perseverance, suggesting more alignment.

For both rigorous expectations and perseverance, the correlations between parents and teacher reports were lower than any other pairs—that is, teacher-student and parent-student—suggesting the least alignment between parents and teachers. These basic correlations indicate how measures of each respondent type move together. For example, they showed that schools with higher student reports tended to have higher teacher reports. However, they did not inform whether the responses align *on average*.

Figure 5. Of the pairs of respondent reports, teacher and parent reports were the least correlated across schools



* Significant at $p < .05$; *** significant at $p < .001$.
# Correlation differed from scale-specific "teacher & parent" (the lowest) correlation by 0.10 or more—the cutoff for a moderate difference between correlations.
Note: The figure shows the pairwise correlation of school-level averages of the scales (described in appendix B) between each pair of respondents.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

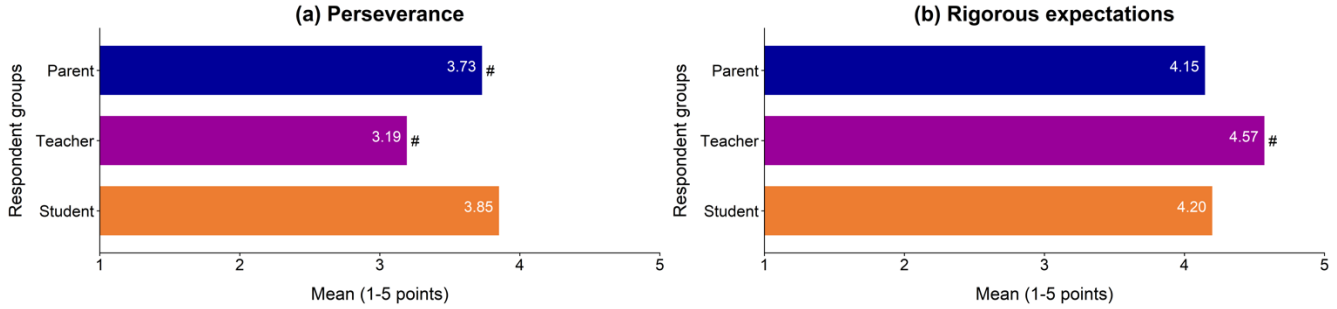
Teachers who responded to the survey provided less favorable ratings of students' perseverance and more favorable ratings of rigorous expectations than did students or parents (figure 6). On average, teacher reports of perseverance were lower than student and parent reports by 0.66 and 0.54 points on a 1- to 5-point scale, respectively. On the other hand, teacher reports of rigorous expectations were higher than student and parent reports by 0.37 and 0.42 points on a 1- to 5-point scale, respectively. Although the student and teacher reports were correlated, this finding indicates that teachers consistently reported different average levels compared to students. The average school-level differences between respondent types were related to the characteristics of

[17] The response rates for the parent survey were substantially lower than the other two, so caution should be used when considering the parent findings.

the schools (including School Transparency and Reporting [STAR] rating and geographic ward), suggesting that alignment varied systematically across schools (see tables C16–C19 in appendix C).

Although it is not possible to explore nonresponse bias for parents, the study team conducted analyses using student and teacher nonresponse weights. Relative to the analyses without weights, the correlations between respondent types changed by at most 0.02 (see figure C6 in appendix C), and the average differences between respondent types changed by at most 0.01 (see figure C7). These estimates suggest that nonresponse bias on the student and teacher surveys likely did not drive the results. However, the study team cannot rule out that nonresponse bias affected the parent survey, so it suggests caution when interpreting those findings.

Figure 6. Average reports on perseverance and rigorous expectations differed across respondent types mainly because teachers responded differently than students and parents



Absolute value of difference relative to students met or exceeded 0.10.
Note: The figure shows the average school-level reports of the scales (described in appendix B) for parents, teachers, and students.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Limitations

Readers should keep one primary limitation in mind as they consider this study's findings: it examined correlational relationships between variables, not causal ones (see appendix B for other limitations). The results suggest which SEL competencies and school experiences were associated with student outcomes. Although it is possible that SEL competencies and school experiences could *cause* those outcomes, it is also possible that the variables were associated because other factors caused better student outcomes as well as better SEL competencies and school experiences. For example, supportive parents could both foster students' SEL competencies and improve students' outcomes by assisting them with schoolwork. Therefore, the findings do not imply that improvements in SEL competencies or school experiences would necessarily improve outcomes.

Implications

The study addressed several research questions that investigated the properties of measures of SEL competencies and school experiences in DCPS. The findings for each research question have implications for DCPS—particularly as they relate to the district's strategic goals (see table A1 in appendix A)—and for other districts with access to similar types of data.

Research question 1. The study findings suggest that middle school and early high school could be especially important times to support students' SEL competencies and improve their experiences at school. During middle school and early high school, students reported the lowest SEL competencies and school experiences. These findings are consistent with other research that has demonstrated that middle school and transitioning to high school can be especially challenging (Benner, 2011; Eccles, 2004; Rudolph et al., 2001). Supporting students before or during these grades could potentially help DCPS achieve its goal of helping all students to feel loved, challenged, and prepared (see table A1 in appendix A). In addition, DCPS could investigate the causes for these declines to better understand how to support students. At the same time, the study cannot rule out that this U-shaped

pattern emerges because the composition of students changed across grades. For example, if less perseverant students dropped out of school after grades 9 and 10, the average perseverance of students who remained enrolled in grades 11 and 12 would increase. The results still address the study's research question, because they suggest which grades DCPS could target to improve the SEL competencies and school experiences for enrolled students—the students whom DCPS can most easily reach through programming or initiatives in schools. Students who have already dropped out would be more challenging to support directly.

Research question 2. Across schools, year-to-year changes in SEL competencies and school experiences differed to a meaningful degree, similar to that of academic measures. These findings could have arisen for at least two reasons. First, differences in practices across schools could have led to differences in year-to-year changes in the measures. In this case, the findings suggest that schools have the potential to improve SEL competencies and school experiences. Second, other factors that were associated with schools could have driven the differences in the year-to-year changes. As a next step, DCPS could explore the extent to which practices differ between schools that are currently associated with higher growth in SEL competencies and school experiences and those with lower levels of growth. If the practices differ systematically, then that would provide additional evidence that school practices are driving the differences, and the higher-change schools could potentially serve as a model for lower-change schools. These findings also suggest that school-level targeting could be effective. In particular, DCPS's student support teams could work with lower-performing schools to develop plans around any promising SEL-related programming. Such efforts could help DCPS achieve its strategic goal that all schools should be highly rated or improving (see table A1 in appendix A).

Research question 3. The study found that students' SEL competencies and school experiences—especially self-management and rigorous expectations—were moderately to substantively related to their later academic outcomes; this finding suggests that improving SEL competencies and school experiences may help DCPS progress toward its strategic goals around high school graduation, re-enrollment, early literacy, and college and career readiness (see table A1 in appendix A). These findings are consistent with the broader literature, which has found correlations of a similar magnitude and that skills related to self-management tend to be the most related to students' future academic outcomes and longer-term outcomes like job performance (Almlund et al., 2011; Poropat, 2009; West et al., 2016). Of the SEL competencies and school experiences, self-management was the most related to students' future outcomes, so DCPS may consider prioritizing strategies designed to promote this competency. However, confirming this priority would require additional research, such as an evaluation of interventions designed to improve SEL competencies and school experiences.

At the same time, measures of SEL competencies and school experiences added little value in classifying whether students were at risk of poor academic outcomes beyond the other predictors. On their own, the SEL competencies and school experiences predicted students' outcomes. However, they did not add predictive power beyond other predictors (demographic variables and prior academic outcomes), because those other predictors may have captured information similar to that captured by measures of SEL competencies and school experiences. For that reason, DCPS may consider focusing on predictive models that use demographic and academic measures as predictors. The measures of SEL competencies and school experiences can still be helpful in supporting students. Once at-risk students are identified, their SEL competencies and school experiences may inform how schools can support them. In addition, SEL competencies and school experiences improved predictions of whether students felt loved, challenged, and prepared, which may be of interest given DCPS's goal around this outcome.

Research question 4. DCPS could consider steps to investigate differences between student, teacher, and parent reports of perseverance and rigorous expectations. For example, compared to students, teachers provided less favorable ratings of students' perseverance and more favorable ratings of rigorous expectations. DCPS could incorporate this type of information in discussions with students and teachers in schools with large apparent differences to better understand their perceptions and inform areas for improvement in school quality.

References

Almlund, M., Duckworth, A., Heckman, J. J., & Kautz, T. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 1–181). Elsevier.

Barbaranelli, C., Caprara, G. V., Rabasca, A., & Pastorelli, C. (2003). A questionnaire for measuring the big five in late childhood. *Personality and Individual Differences*, *34*(4), 645–664.

Benner, A. D. (2011). The transition to high school: Current knowledge, future directions. *Educational Psychology Review*, *23*(3): 299–328.

District of Columbia Public Schools (DCPS). (2017). *DCPS strategic plan – A capital commitment 2017–2022*. https://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/DCPS%20Strategic%20Plan%20-%20A%20Capital%20Commitment%202017-2022-English_0.pdf

District of Columbia Public Schools (DCPS). (2019). *DCPS 2019 Panorama survey results*. https://dcps.dc.gov/sites/default/files/dc/sites/dcps/page_content/attachments/2019-Panorama-Survey-Report-FINAL.pdf

Eccles, J. S. (2004). Schools, academic motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 125–153). John Wiley and Sons.

Elango, S., Hojman, A., García, J. L., & Heckman, J. J. (2016). Early childhood education. In R. Moffitt (Ed.), *Means-tested transfer programs in the United States II* (pp. 235–298). University of Chicago Press.

Elliott, S. N., Anthony, C. J., DiPerna, J. C., Lei, P.-W., & Gresham, F. M. (2020). *SSIS SEL Brief + Mental Health Scales expanded user guide & technical manual*. SAIL CoLab, LLC.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics, 19*(4), 451–464.

Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success* (OECD Education Working Papers No. 110). OECD Publishing.

O'Conner, R., De Feyter, J., Carr, A., Luo, J. L., & Romm, H. (2017). *A review of the literature on social and emotional learning for students ages 3–8: Characteristics of effective social and emotional learning programs* (Part 1 of 4) (REL 2017–245). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. https://eric.ed.gov/?q=ED572721&id=ED572721

Panorama Education. (n.d.) Measure and understand social-emotional learning. https://www.panoramaed.com/social-emotional-learning-sel

Poropat, A. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338.

Rudolph, K. D., Lambert, S. F., Clark, A. G., & Kurlakowsky, K. D. (2001). Negotiating the transition to middle school: The role of self-regulatory processes. *Child Development*, *72*(3), 929–946.

Schochet, P. Z., & Chiang, H. S. (2010). Error rates in measuring teacher and school performance based on student test score gains (NCEE 2010-4004). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

SRI International. (2018). Promoting grit, tenacity, and perseverance: Critical factors for success in the 21st century. https://www.sri.com/publication/promoting-grit-tenacity-and-perseverance-critical-factors-for-success-in-the-21st-century/

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 students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis*, *38*(1), 148–170.

West, M. R., Pier, L., Fricke, H., Loeb, S., Meyer, R. H., & Rice, A. B. (2018). *Trends in student social-emotional learning: Evidence from the CORE districts.* Policy Analysis for California Education.

West, M. R., Pier, L. Fricke, H., Hough, H., Loeb, S., Meyer, R. H., & Rice, A. B. (2020). Trends in student social-emotional learning: Evidence from the first large-scale panel student survey. *Educational Evaluation and Policy Analysis*, *42*(2), 279–303.

REL 2021–114

August 2021

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-17-C-0006 by Regional Educational Laboratory Mid-Atlantic administered by Mathematica. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Kautz, T., Feeney, K., Chiang, H., Lauffer, S., Bartlett, M., & Tilley, C. (2021). *Using a survey of social and emotional learning and school climate to inform decisionmaking* (REL 2021–114). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. http://ies.ed.gov/ncee/edlabs

This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

Using a survey of social and emotional learning and school climate to inform decisionmaking

Appendix A. About the study

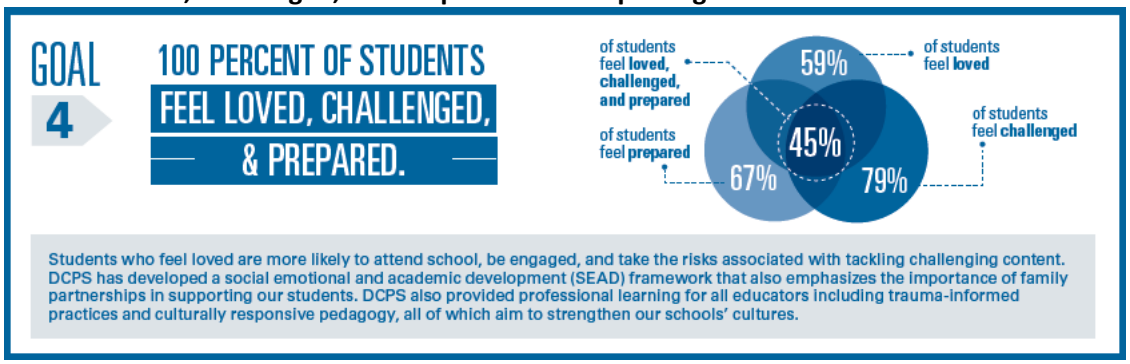
The District of Columbia Public Schools (DCPS) needed timely information on its social and emotional learning (SEL) data in order to meet its 2017–22 Strategic Plan goals. The Strategic Plan is an effort that DCPS uses to hold itself accountable to the public. In its 2017 Strategic Plan, DCPS set six concrete and ambitious goals to "[become] a district of both excellence and equity" (DCPS, 2017). The Strategic Plan highlights SEL competencies and school climate as key components in its strategic priority to "educate the whole child" (DCPS, 2017). All six goals reflect this priority, either explicitly or implicitly. For example, DCPS has set an explicit goal that by 2022, 100 percent of students will feel "loved, challenged, and prepared," as measured by an index based on measures of SEL competencies and school climate (DCPS, 2017). Other goals—such as improving college and career readiness and re-enrollment rates—are implicitly related to SEL competencies in that they could be furthered by better identifying at-risk students and boosting SEL competencies.

DCPS requested the Regional Educational Laboratory (REL) Mid-Atlantic's support in providing information about how to achieve these goals or make substantial progress toward them by 2022. These analyses of SEL and climate measures and related administrative data (for example, absences and credits earned) provide DCPS this information in order to reach its time-sensitive Strategic Plan goals.

This study builds on DCPS's recent work to prioritize and measure SEL competencies and climate perceptions since it began administering the Panorama Education survey in spring 2018. It examines the properties of SEL and climate measures based on these surveys to inform progress toward DCPS's six Strategic Plan goals. Supported through a 2018/19 REL Mid-Atlantic coaching activity, DCPS has successfully developed and validated measures of SEL competencies based on the SEL survey. Each measure captures a different SEL competency (for example, perseverance) or aspect of school climate (for example, rigorous expectations) and is based on averaging scores of a group of items designed to measure that construct. The student measures span four SEL competencies (perseverance, self-management, self-efficacy, and social awareness) and three measures of a school's climate that could boost SEL competencies (rigorous expectations, sense of belonging, and student satisfaction). As part of the earlier REL Mid-Atlantic coaching activity, it was demonstrated that the measures met standard criteria for reliability and validity (Kautz et al., 2019). In addition, the study team used five of these measures to develop the Loved, Challenged, and Prepared Index that DCPS is now using to track and publicly report progress toward its ambitious goal that 100 percent of students will feel "loved, challenged, and prepared" by 2022 (figure A1).[18]

[18] Using theoretical and empirical evidence, the measures were assigned to each component as follows: sense of belonging capturing loved; rigorous expectations capturing challenged; and perseverance, self-management, and self-efficacy capturing prepared.

Figure A1. DCPS's Loved, Challenged, and Prepared Index reporting



DCPS is District of Columbia Public Schools.
Source: DCPS, 2018

While DCPS has made much progress, it still has a long way to go to achieve these goals (table A1). For example, to meet Goal 3, DCPS needs to boost its graduation rate by over 15 percentage points, a substantial amount.

Table A1. How this study addresses each Strategic Plan goal

Goal	Year 1 reported status[a]	How this study could contribute to this goal
1. **Double the percent of students who are college and career ready** as measured by proficiency on the PARCC. This goal represents an increase from 31.9 percent to 63.8 percent proficient for ELA and 27.4 percent to 54.8 percent proficient for math	• 35.1 percent for proficient for ELA • 30.5 percent proficient for math	• Use DCPS's data to suggest which SEL competencies or aspects of school climate might be most promising to target to boost college and career readiness.
2. **100 percent of K–2 students reading on or above grade level**	• 65 percent are reading on or above grade level	• Use DCPS's data to suggest which SEL competencies or aspects of school climate might be most promising to target to improve early reading skills.
3. **85 percent of students graduate within four years and 90 percent graduate within five years**	• 68.6 percent graduate in four years • 75.5 percent graduate in five years	• Provide DCPS with new ways to identify students at risk of dropping out or not transitioning between grades. • Use DCPS's data to suggest which SEL competencies or aspects of school climate to focus on for individual students.
4. **100 percent of students feel loved, challenged, and prepared**	• 45 percent of students feel loved, challenged, and prepared	• Provide information to help DCPS identify the types of students who could benefit from additional support (such as by age, grade, and demographic group). • Inform whether it might be fruitful to focus on particular schools or particular types of schools.
5. **100 percent of schools highly rated or improving** based on the DC School Transparency and Reporting Framework, which incorporates test scores, advanced classes enrollment rate, adjusted cohort graduation rate, in-seat attendance rate, re-enrollment	• Not available	• Provide information to help identify which SEL competencies or aspects of school climate that schools could target to improve school ratings.

Goal	Year 1 reported status[a]	How this study could contribute to this goal
6. **90 percent of students re-enroll and DCPS serves 54,000 students**	• 84 percent of students re-enrolled • 49,103 enrolled in DCPS	• Provide DCPS with new ways to identify students at risk of not re-enrolling. • Provide DCPS with information on which SEL competencies or aspects of school climate to focus on for improving individual students' re-enrollment.

DCPS is District of Columbia Public Schools. ELA is English language arts. PARCC is Partnership for Assessment of Readiness for College and Career. SEL is social and emotional learning.

[a] Based on DCPS's one-year capital commitment update (DCPS, 2018).

This study provides a range of information that DCPS can use to better achieve its six strategic goals, with a particular focus on Goal 4, which directly relates to SEL competencies and school climate. In particular, the study provides key information on the types of students who are not feeling loved, challenged, and prepared as well as the types of schools they attend. Such findings may suggest how DCPS can best target its investments to make the biggest difference to progress toward Goal 4, which is especially pertinent because investment in SEL across the district is on the rise. In addition to supporting a district-level SEL team, the district has hired school-based SEL leads, who work with principals to create school-level SEL and climate goals and, in some cases, have begun to administer additional short surveys around these topics at the beginning, middle, and end of the school year. Because these grassroots efforts are not yet standardized across schools, a better understanding of its Panorama SEL data may allow the district to create well-informed, universal SEL and climate practices or programming to attain its goals.

While Goals 1, 2, 3, 5, and 6 are longer-term and less explicitly tied to SEL and school climate, the study will provide information about how boosting SEL competencies and improving school climate could be an avenue for achieving them.

References

District of Columbia Public Schools (DCPS). (2017). *DCPS strategic plan – A capital commitment 2017–2022.* https://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/DCPS%20Strategic%20Plan%20-%20A%20Capital%20Commitment%202017-2022-English_0.pdf

District of Columbia Public Schools (DCPS). (2018). *A capital commitment 2017–2022 – year 1 update.* https://dcps.dc.gov/publication/capital-commitment-2017-2022-year-1-update

Kautz, T., Feeney, K., & Glazerman, S. (2019). *Developing and validating socio-emotional learning measures in DC.* Accountability alliance meeting, Regional Educational Laboratory Mid-Atlantic. https://ies.ed.gov/ncee/edlabs/regions/midatlantic/app/Docs/Events/Accountability_in_ESSA_era_alliance_meeting_April_2019.pdf

Appendix B. Methods

This appendix describes the data, samples, weights, and analysis methods for the Regional Educational Laboratory Mid-Atlantic study to explore the properties and uses of social and emotional learning (SEL) survey data to inform evaluation and track progress toward District of Columbia Public School (DCPS) goals.

Data

The key data sources for this study are DCPS's Panorama Education (Panorama) surveys and administrative records.

Panorama surveys. The Panorama survey is administered annually to students, teachers, and parents. All students between grades 3 and 12 who attended at least one day of school are eligible to complete the survey, while all parents and teachers are eligible. If a parent has multiple children in a single school, the parent is asked to complete one survey with the oldest child in mind (DCPS, 2020). If the parent has multiple children enrolled in different schools, the parent is asked to complete one survey for each school. The student survey assesses four SEL competencies—perseverance, self-management, self-efficacy, and social awareness—and three aspects of school climate—rigorous expectations, sense of belonging, and student satisfaction. The number of items in each topic ranges from four to eight, and the version for grades 3–5 includes fewer items for perseverance and sense of belonging (table B1). DCPS also surveys teachers and parents to assess their perceptions of students' SEL competencies and personal experiences of school climate that are related to perseverance and rigorous expectations. The teacher surveys include questions about these selected competencies of their students as a group, as well as their own engagement and professional learning about SEL. The parent surveys include questions about their child's SEL competencies, as well as questions about their own experience with the school and individual teachers. When referring to individual respondents' reports, this study uses the term school experiences. Averaging the measures of school experiences across respondents at a school provides a measure of a school's climate.

Table B1. Number of items in DCPS's customized Panorama Education survey

Respondent type	Perseverance	Self-management	Self-efficacy	Social awareness	Rigorous expectations	Sense of belonging	Student satisfaction
Student							
Grades 3–5	4	5	5	8	5	4	8
Grades 6–12	5	5	5	8	5	5	8
Parent	6	na	na	na	3	na	na
Teacher	5	na	na	na	5	na	na

DCPS is District of Columbia Public Schools. na is not applicable.

For each SEL competency and school climate measure, the associated scales are calculated by averaging across numerical values that correspond to each possible response category. Each response category is assigned a value between 1 and 5, so the resulting scales range from 1 to 5 points.[19] Following DCPS's methodology, this study counted the scale as non-missing if a respondent completed at least two items in the scale.

The administration of the Panorama survey was directly impacted by the outbreak of the COVID-19 pandemic. The survey, which is typically administered online to students in school in early March, was only available online for students to complete at home once students could no longer attend school in person, resulting in a lower response rate for students (51 percent, which is down nearly 20 percent from the previous year). For all

[19] The student satisfaction items have four response categories. When reporting findings, the study team rescaled the student satisfaction scale to range from 1 to 5, so it was comparable to the other scales. To do so, the study team rescaled the values linearly so that a 2 on the original scale mapped to a 2.33 on the updated scale, a 3 mapped to a 3.66, and a 4 mapped to a 5.

respondents, the survey was likely not a top priority given the pandemic's upheaval to normal life, and those who did respond might have responded differently as a result. For these reasons, our main analyses did not include the 2019/20 survey data.

Over the two school years included in the study (2017/18–2018/19), there were 39,791 unique students enrolled and 30,462 who responded to the Panorama survey (table B2). Analyses that include parent and teacher samples were also run over only two years of data (2017/18–2018/19). Over this period, there were 4,273 unique teachers who responded and 12,216 responses from parents on the measures used in this study.

Table B2. Sample sizes by respondent type and school year

Sample description	School year(s)	Sample size
Unique students who were eligible for the survey	2017/18–2018/19	39,791
	2017/18	31,452
	2018/19	31,926
Unique student respondents for the survey	2017/18–2018/19	30,462
	2017/18	21,385
	2018/19	22,208
Unique teacher respondents for the survey measures used in this study	2017/18–2018/19	4,273
	2017/18	3,109
	2018/19	3,130
Parent responses to the survey measures used in this study[a]	2017/18–2018/19	12,216
	2017/18	5,734
	2018/19	6,482

[a] These calculations do not represent unique parent respondents, because they could include multiple responses from the same person.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Administrative data. The student administrative data files included information for all DCPS students from fall 2017 through spring 2020. The data included demographic information, such as gender, special education status, English learner status, grade level, race/ethnicity, and whether they are at risk (per DCPS's classification; table B3). They also included academic measures, such as absences and in-seat attendance rates, suspensions, grade progression, enrollment status, math and English language arts (ELA) summative assessment proficiencies on the Partnership for Assessment of Readiness for College and Careers (PARCC) or Multi-State Alternate Assessment (MSAA), predicted college readiness based on SAT or Preliminary SAT (PSAT) scores, and credits earned.[20] In addition, the data files included teacher demographic information (such as gender, race/ethnicity, and employment duration) and school characteristics (such as DC School Transparency and Reporting [STAR] rating and geographic ward).

Similar to the Panorama survey data, the outbreak of the COVID-19 pandemic affected the availability of administrative data during the 2019/20 school year. Certain annual standardized tests were canceled in spring 2020, including the PARCC exams and some spring administrations of the SAT and PSAT.

[20] MSAA scores were only received in school year 2018/19.

Table B3. Student characteristics and outcomes used in the study

Characteristic	Description
Demographic variables	
Female	Whether a student was female
Age	Age of a student as of September 30th for that academic year
Grade level	Which grade (3–12) a student was in at the time of enrollment for that academic year
Survey year	Which academic year data was from (2017/18 or 2018/19)
Race/ethnicity	Whether a student was • Black (non-Hispanic) • Hispanic • White (non-Hispanic) • Other
At-risk status	Whether a student was flagged as at risk. "In the District of Columbia, at-risk is defined as a student who possesses one of the following characteristics at any point during the given school year: eligibility for Temporary Assistance for Needy Families (TANF), eligibility for Supplemental Nutrition Assistance Program (SNAP), identification as homeless by the student's school or other community partners, under the care of the Child and Family Services Agency (CFSA, also known as foster care), and/or overage (high school only). A high school student is overage if he or she is at least one year older than the expected age for their grade" (District of Columbia Office of the State Superintendent of Education, 2020).
Special education status	Whether a student received special education services
English learner status	Whether a student was an English learner
Ward	In which of eight geographic areas, called a ward, a student attended school
STAR rating	To which of five rating scores a student's school was assigned. "The School Transparency and Reporting (STAR) Framework is the accountability framework for public schools in the District of Columbia. The STAR Framework uses common measures of performance across schools and is comprised of multiple data points from multiple data sources. Schools receive a summative STAR Rating ranging from 1 to 5 stars, with 5 being the highest" (District of Columbia Office of the State Superintendent of Education, 2020).
Academic behavior measures	
In-seat attendance	The number of days a student attended school during the school year divided by the number of days they were enrolled (at the school where students spent the majority of time during the school year)
Chronically absent	Whether a student was chronically absent (absent for 10 percent or more of the school year)
Number of suspensions	The number of times a student was suspended during the school year
Suspended	Whether a student was ever suspended during the school year
Progressed successfully	Whether a student who was enrolled in 2017/18 progressed as expected in 2018/19 and 2019/20. For example, if the student was in grade 3 in 2017/18, whether they were in grade 4 in 2018/19 and grade 5 in 2019/20. For students in grade 12, this would be whether they graduated. Progression also included students who skipped a grade (for example, a student in grade 3 in 2017/18 and grade 5 in 2018/19).
Number of credits behind	The number of credits high school students were behind for a given grade level
Successful re-enrollment	Whether a student who was enrolled in 2017/18 remained enrolled through later school years
Red status for grade progression	Whether a student was flagged as being 4.25+ credits behind in order to graduate on time, assuming students earn 6 credits per year in high school

Table B3. Student characteristics and outcomes used in the study (continued)

Characteristic	Description
Academic proficiency measures	
Number of AP credits earned	The number of AP credits a student completed/earned during the school year
Earned AP credit	Whether a student completed/earned any AP credits during the school year
Proficient/college ready in math and ELA	In 2017/18, whether a grade 3–10 student was proficient in math/ELA based on their PARCC score (a score of 4+ on a scale of 1 to 5) and whether a grade 11–12 student was college ready in math/reading/writing based on their SAT score (a score of 530+ in math and 480+ in EBRW). In 2018/19, whether a grade 3–10 student was proficient in math/ELA based on their PARCC score or whether a grade 3–8 student was proficient in math/ELA based on their MSAA score (a score of 3+ on a scale of 1 to 4) and whether a grade 11–12 student was college ready in math/reading/writing based on their SAT score. Students in applicable grades in these years took at most either PARCC or MSAA, not both.
Proficient on the PARCC in math	Whether a student was proficient in math based on their PARCC score (a score of 4+ on a scale of 1 to 5)
Proficient on the PARCC in ELA	Whether a student was proficient in ELA based on their PARCC score (a score of 4+ on a scale of 1 to 5)
College ready in math	Whether a student was college ready in math based on their best math SAT score (530+) or their best math PSAT score (PSAT college-ready score threshold is based on grade level)
College ready in ELA	Whether a student was college ready in reading/writing based on their best EBRW SAT score (480+) or their best EBRW PSAT score (PSAT college-ready score threshold is based on grade level)
Math and ELA achievement	The average math and ELA achievement tests score of the student
Outcomes based on SEL competencies and school experiences	
Student Loved, Challenged, and Prepared Index	Whether or not a student felt loved, challenged, and prepared, based on the student's reports of SEL competencies and school experiences. In the index, "loved" is captured by items on sense of belonging, "challenged" is captured by items on rigorous expectations, and "prepared" is captured by items on perseverance, self-management, and self-efficacy. The index is formed in two steps. First, for each component, students are assigned a value of 1 if their average score on the scales associated with that component exceeds 3.5 on 1- to 5-point scale, and they are assigned a value of 0 otherwise. For example, if a student's score on the sense of belonging scale exceeded 3.5, they would be assigned a value of 1 on the loved component, indicating that they feel loved. Second, if they are assigned a value of 1 for all three components, then they are assigned a value of 1 for the Student Loved, Challenged, and Prepared Index. If they are assigned a value of 0 for any component, they are assigned a value of 0 for the Index. This student outcome variable was developed to measure progress toward DCPS's strategic goal of having 100 percent of students feel loved, challenged, and prepared by 2022.

AP is Advanced Placement. DCPS is District of Columbia Public Schools. EBRW is Evidence-Based Reading and Writing. ELA is and English language arts. MSAA is Multi-State Alternate Assessment. PARCC is Partnership for Assessment of Readiness for College and Careers. PSAT is Preliminary SAT. SEL is social and emotional learning. STAR is School Transparency and Reporting.

Data acquisition

DCPS generated a unique district identification number (StudentID) for all enrolled students and provided the study team with the data associated with each student for 2017/18–2019/20, identified by StudentID. In addition, teacher and parent SEL files had unique respondent identification numbers. The data did not include student, teacher, or parent names, addresses, or social security numbers, but the study team took steps to protect the data that included StudentIDs.

Sample

The sample sizes differed by analysis (table B4). In some cases, the sample sizes counted the same students once in each observed year over multiple years, such as the analyses of students' SEL competencies and school experiences for research question 1. Therefore, students in these analyses could be counted multiple times if they appeared in multiple years. In other cases, the analyses required that individual students have data in multiple

years, such as the analyses of the change in SEL competencies and school experiences across years for research question 2.

Table B4. Sample sizes by research question and analysis sample for the main analyses

Research question (RQ)	Analysis sample	Sample size
RQ1. Differences in students' SEL competencies and school experiences across grades	Students who were eligible for the Panorama survey and completed it in 2017/18 or 2018/19	43,578 student-year observations 6,222 teacher-year observations
RQ2. Differences in year-to-year changes in measures across schools	Students who were eligible for the Panorama survey and completed it in 2017/18 and 2018/19	13,131 students
RQ3. Correlations between predictors and outcomes and classification accuracy of predictors	Students who were eligible for the Panorama survey and completed it in 2017/18 and had non-missing data on each predictor	18,152 students
RQ3. Correlations between early literacy and ELA achievement	Students who were in grade 3 in 2017/18 and completed the Panorama survey in 2017/18 or were in grade 3 in 2018/19 and completed the Panorama survey in 2018/19	6,665 student-year observations
RQ3. Correlations between grade 8 measures and grade 9 measures	Students who were in grade 8 in 2017/18, eligible for the Panorama survey and completed it in 2017/18, and had non-missing data on each predictor	1,394 students
RQ4. Correlations between school-level averages of student, parent, and teacher reports	Students, parents, and teachers who were eligible for the Panorama survey and completed at least one of the measures used in this study in 2017/18 or 2018/19; schools that had at least one student complete the Panorama survey in 2017/18 or 2018/19	43,116 student-year observations, 12,216 parent-year observations, 6,222 teacher-year observations, 233 schools
RQ4. Differences between student, parent, and teacher reports	Students, parents, and teachers who were eligible for the Panorama survey and completed at least one of the measures used in this study in 2017/18 or 2018/19; schools that had at least one student complete the Panorama survey in 2017/18 or 2018/19	43,116 student-year observations, 12,216 parent-year observations, 6,222 teacher-year observations, 233 schools

ELA is English language arts. SEL is social and emotional learning.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Analysis methods

Accounting for survey nonresponse.

The study team explored and accounted for nonresponse bias for the SEL competencies and school experiences in the student and teacher surveys. The usefulness of a survey hinges on obtaining responses from a representative set of respondents; if only a select type of person responds, then the resulting findings could paint a misleading picture of the population of interest. The study team explored response rates and the potential for nonresponse bias, and it constructed nonresponse weights to mitigate that bias.

Response rates. The study team calculated the following rates for each analysis sample and survey measure (tables B5 and B6):

- Unit response rate, which is the number of respondents who took the survey out of those who were eligible to take the survey.

- Item response rate, which is the number of respondents who had non-missing values for a scale out of those who took the survey.
- Overall response rate, which is the number of respondents who had non-missing values for a scale out of those who were eligible to take the survey.[21]

For students, the response rates were calculated for six different samples, which included three samples based on students who were eligible to take the survey in individual years (for example, 2017/18) and three samples based on whether students were eligible to take the survey in two years (for example, both 2017/18 and 2018/19). For the samples based on individual years, the overall response rates ranged from 50.7 percent to 69.1 percent. For samples based on two years, the overall response rates ranged from 37.3 percent to 55.0 percent. The results revealed that unit nonresponse—rather than item nonresponse—drove the overall response rates for each scale.

For teachers, the response rates were calculated for two different samples: teachers who were eligible to take the survey in 2017/18 and teachers who were eligible to take it in 2018/19. The overall response rates ranged from 59.7 to 77.1. As with the student survey, unit nonresponse drove the overall response rates for each scale.

Nonresponse bias analysis. Because the overall response rates were less than 85 percent for each of our samples, a conventional standard for nonresponse bias to be considered negligible, the study team conducted a nonresponse bias analysis. The team compared the characteristics of the students and teachers who responded to the survey to the full sample of those who were eligible. For students, the study has administrative records on a number of other characteristics, including their grade level, gender, English learner status, special education status, and in-seat attendance, as well as the number of suspensions they received and whether they are at risk (as defined by the DCPS indicator). For teachers, the analyses included information on their gender, race/ethnicity, and level of experience, as well as information on their school—including the type of school, its STAR rating, the geographical ward, the racial/ethnic composition of its students, the average math and ELA achievement tests scores of the students, the average number of suspensions per student, and the average in-seat attendance for students. These variables were all missing in less than 5 percent of the sample cases.

A nonresponse bias analysis requires having at least one characteristic that is strongly related to each survey measure. To assess the relationship between the survey measures and the characteristics, the study team calculated the correlations between them. Because all the individual characteristics had relatively low correlations with the survey measures, the study team constructed a composite variable of characteristics that would be more highly correlated with the survey measures. Compared to the individual measures, the composite variables were more correlated with the survey measures, because the composite variables summarized information from multiple individual measures, which can reduce measurement error. Specifically, to form the composite variable, the study team conducted an ordinary least squares regression of each of the survey measures on the group of student and teacher characteristics. The study team then formed the composite variable as the predicted value of the survey measure for each student in the sample. For the samples based on individual years of data, the correlations between the student survey measures and the composite variable ranged from 0.22 to 0.39 (tables B7–B13). The study team deemed that these correlations were strong because they approached the rule-of-thumb cutoff for a strong correlation of 0.25. For the samples based on two years of data, the correlations between changes in the student survey measures and the composite variable ranged from 0.07 to 0.28 (tables B14–B20). Although these correlations were not as strong, the study team is not aware of evidence on variables that have a stronger correlation with changes in measures of SEL competencies and school experiences. The correlations between the teacher survey measures and the composite variable ranged from 0.23 to 0.40 (tables B21–B22).

[21] To be considered not missing, a respondent needed to answer at least two items for the scale.

Because they approached the 0.25 cutoff for a strong correlation, the study team deemed that these correlations were strong.

For the composite variable and each of these characteristics, the study team calculated the difference in standard deviation units between respondents to the survey and the full samples of students and teachers who were eligible (tables B7–B22). At least one difference exceeded 0.05 standard deviations for each survey measure, providing evidence of nonresponse bias. Notably, the tables reveal that, because students in higher grades responded less frequently, they were underrepresented in the survey sample—compared to those in lower grades—for each survey measure.

Nonresponse response weights. Given the evidence of nonresponse bias, the study team constructed nonresponse weights for students and teachers. Nonresponse weights help to ensure that the results are statistically representative, reflecting the composition of students in DCPS by giving higher weight to the types of respondents who are less likely to respond. To form nonresponse weights, the team adopted the following steps:

1. Using a probit model, the study team estimated the probability of completing the survey as a function of the student characteristics listed in tables B7–B22.
2. Using estimates from each of the probit models, the study team calculated the propensity of being in the index sample for each respondent in the sample ($Propensity\_i$). Next, the study team calculated the deciles of the distribution of propensities and assigned each respondent the average value within their decile ($Propensity\_decile\_i$).
3. Finally, the study team formed nonresponse weights by taking the inverse of the average decile value for each respondent ($1/Propensity\_decile\_i$).

To assess the effectiveness of the weights, the study team compared the difference in the composite of covariates between the survey sample and the original sample with and without nonresponse weights (table B23). For both the student and teacher surveys, the absolute difference with weights was less than 0.05 standard deviations for the individual years on which the study focuses (2017/18 or 2018/19)—a rule of thumb for an acceptable difference. For the analyses that the study focused on (that involved students who were eligible in two years: 2017/18 and 2018/19), all but one absolute difference was 0.07 standard deviations or less. The remaining difference was 0.09 standard deviations. For the samples the study did not use in the main analyses, the differences tended to be greater.

Table B5. Response rates (student survey)

Type of data the study attempted to collect	2017/18	2018/19	2019/20	2017/18 and 2018/19	2018/19 and 2019/20	2017/18 and 2019/20
Perseverance						
Unit response rate	68.0	69.6	51.3	55.7	41.8	39.4
Item response rate	95.7	98.5	99.1	94.3	97.6	94.6
Overall response rate	65.1	68.5	50.9	52.5	40.8	37.3
Self-management						
Unit response rate	68.0	69.6	51.3	55.7	41.8	39.4
Item response rate	96.4	97.8	98.7	94.6	96.9	95.6
Overall response rate	65.6	68.0	50.7	52.6	40.5	37.6
Self-efficacy						
Unit response rate	68.0	69.6	51.3	55.7	41.8	39.4
Item response rate	96.6	97.8	99.0	94.8	97.0	95.9
Overall response rate	65.7	68.0	50.8	52.8	40.5	37.8
Social awareness						
Unit response rate	68.0	69.6	51.3	55.7	41.8	39.4
Item response rate	96.5	98.0	98.9	94.7	97.2	95.5
Overall response rate	65.6	68.2	50.7	52.7	40.6	37.6
Rigorous expectations						
Unit response rate	68.0	69.6	51.3	55.7	41.8	39.4
Item response rate	97.5	99.2	99.5	96.8	98.7	97.3
Overall response rate	66.3	69.0	51.1	53.9	41.2	38.3
Student satisfaction						
Unit response rate	68.0	69.6	51.3	55.7	41.8	39.4
Item response rate	99.4	99.4	99.8	98.9	99.3	99.3
Overall response rate	67.6	69.1	51.2	55.0	41.5	39.1
Sense of belonging						
Unit response rate	68.0	69.6	51.3	55.7	41.8	39.4
Item response rate	98.9	99.3	99.5	98.2	98.9	98.3
Overall response rate	67.3	69.1	51.1	54.7	41.3	38.7

Note: The table shows the unit, item, and overall response rates to the student survey for each of the seven survey scales. The item response rate is the percentage of observations for which each scale was missing among those who responded to the survey.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B6. Response rates (teacher survey)

Type of data the study attempted to collect	2017/18	2018/19
Perseverance		
Unit response rate	60.1	77.3
Item response rate	99.3	99.7
Overall response rate	59.7	77.1
Rigorous expectations		
Unit response rate	60.1	77.3
Item response rate	99.3	99.8
Overall response rate	59.7	77.1

Note: The table shows the unit, item, and overall response rates to the teacher survey for each of the two survey scales used in the study. The item response rate is the percentage of observations for which each scale was missing among those who responded to the survey.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table B7. Exploration of nonresponse bias for perseverance (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with perseverance scale		
	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20
Female	0.50	0.50	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.03	0.03	0.05	0.02	0.00
Grade level												
3	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.34)	0.13 (0.33)	0.09	0.07	0.13	0.10	0.10	0.10
4	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.33)	0.12 (0.33)	0.09	0.08	0.13	0.08	0.09	0.07
5	0.15	0.15	0.16	0.12 (0.32)	0.12 (0.32)	0.11 (0.32)	0.09	0.08	0.13	0.05	0.05	0.06
6	0.08	0.10	0.11	0.08 (0.27)	0.09 (0.29)	0.09 (0.29)	0.00	0.02	0.06	0.00	-0.02	-0.04
7	0.09	0.09	0.11	0.08 (0.27)	0.08 (0.27)	0.09 (0.29)	0.03	0.04	0.07	-0.05	-0.07	-0.08
8	0.07	0.09	0.09	0.07 (0.26)	0.08 (0.27)	0.08 (0.28)	0.00	0.03	0.04	-0.05	-0.08	-0.08
9	0.08	0.07	0.06	0.10 (0.31)	0.10 (0.31)	0.12 (0.32)	-0.09	-0.10	-0.17	-0.08	-0.05	-0.06
10	0.07	0.06	0.04	0.10 (0.30)	0.09 (0.29)	0.09 (0.28)	-0.09	-0.10	-0.17	-0.07	-0.06	-0.04
11	0.07	0.07	0.04	0.09 (0.28)	0.08 (0.28)	0.08 (0.26)	-0.06	-0.07	-0.13	-0.05	-0.04	-0.02
12	0.07	0.06	0.04	0.09 (0.29)	0.09 (0.28)	0.08 (0.27)	-0.10	-0.09	-0.16	-0.01	-0.01	-0.01
English learner student	0.12	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.14 (0.34)	0.01	-0.02	-0.03	0.01	0.03	0.02
Special education student	0.14	0.14	0.14	0.16 (0.37)	0.17 (0.37)	0.17 (0.37)	-0.06	-0.07	-0.07	-0.03	-0.02	-0.02
Fraction of days absent	0.09	0.09	0.07	0.15 (0.20)	0.15 (0.21)	0.13 (0.19)	-0.27	-0.27	-0.32	-0.08	-0.06	-0.05
Number of suspensions	0.14	0.16	0.07	0.18 (0.65)	0.20 (0.74)	0.09 (0.43)	-0.05	-0.05	-0.07	-0.06	-0.05	-0.03
At-risk status	0.42	0.43	0.42	0.42 (0.49)	0.48 (0.50)	0.48 (0.50)	0.00	-0.10	-0.13	0.08	0.08	0.09
Composite of covariates	**3.78**	**3.78**	**3.77**	**3.75 (0.19)**	**3.75 (0.19)**	**3.73 (0.18)**	**0.19**	**0.13**	**0.22**	**0.24**	**0.25**	**0.24**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B8. Exploration of nonresponse bias for self-management scale (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with self-management scale		
	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20
Female	0.50	0.50	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.03	0.04	0.10	0.08	0.06
Grade level												
3	0.16	0.16	0.17	0.13 (0.34)	0.13 (0.34)	0.13 (0.33)	0.10	0.07	0.14	0.02	0.04	0.03
4	0.16	0.16	0.17	0.13 (0.34)	0.13 (0.33)	0.12 (0.33)	0.09	0.09	0.13	0.03	0.04	0.04
5	0.15	0.15	0.16	0.12 (0.32)	0.12 (0.32)	0.11 (0.32)	0.10	0.09	0.13	0.00	0.00	0.02
6	0.08	0.10	0.11	0.08 (0.27)	0.09 (0.29)	0.09 (0.29)	0.00	0.02	0.06	0.00	0.00	-0.02
7	0.09	0.09	0.11	0.08 (0.27)	0.08 (0.27)	0.09 (0.29)	0.02	0.03	0.07	-0.02	-0.06	-0.05
8	0.07	0.09	0.09	0.07 (0.26)	0.08 (0.27)	0.08 (0.28)	0.00	0.02	0.04	-0.03	-0.03	-0.03
9	0.08	0.07	0.06	0.10 (0.31)	0.10 (0.31)	0.12 (0.32)	-0.09	-0.10	-0.17	-0.03	-0.02	-0.02
10	0.07	0.06	0.04	0.10 (0.30)	0.09 (0.29)	0.09 (0.28)	-0.09	-0.10	-0.17	-0.03	-0.02	0.00
11	0.07	0.06	0.04	0.09 (0.28)	0.08 (0.28)	0.08 (0.26)	-0.06	-0.07	-0.13	0.01	0.00	0.01
12	0.07	0.06	0.04	0.09 (0.29)	0.09 (0.28)	0.08 (0.27)	-0.10	-0.09	-0.16	0.04	0.02	0.00
English learner student	0.12	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.14 (0.34)	0.01	-0.02	-0.03	-0.02	-0.03	-0.04
Special education student	0.14	0.14	0.14	0.16 (0.37)	0.17 (0.37)	0.17 (0.37)	-0.06	-0.07	-0.07	-0.10	-0.11	-0.10
Fraction of days absent	0.09	0.09	0.07	0.15 (0.20)	0.15 (0.21)	0.13 (0.19)	-0.27	-0.27	-0.32	-0.11	-0.12	-0.09
Number of suspensions	0.14	0.16	0.06	0.18 (0.65)	0.20 (0.74)	0.09 (0.43)	-0.06	-0.06	-0.07	-0.14	-0.14	-0.10
At-risk status	0.42	0.43	0.42	0.42 (0.49)	0.48 (0.50)	0.48 (0.50)	0.00	-0.10	-0.13	-0.09	-0.10	-0.11
Composite of covariates	**3.77**	**3.80**	**3.78**	**3.74 (0.20)**	**3.76 (0.20)**	**3.74 (0.18)**	**0.18**	**0.19**	**0.23**	**0.24**	**0.24**	**0.22**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B9. Exploration of nonresponse bias for self-efficacy (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with self-efficacy scale		
	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20
Female	0.50	0.50	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.03	0.04	-0.02	-0.04	-0.06
Grade level												
3	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.34)	0.13 (0.33)	0.09	0.07	0.14	0.11	0.12	0.10
4	0.16	0.16	0.17	0.13 (0.34)	0.13 (0.33)	0.12 (0.33)	0.09	0.08	0.13	0.12	0.12	0.10
5	0.15	0.15	0.16	0.12 (0.32)	0.12 (0.32)	0.11 (0.32)	0.09	0.09	0.14	0.10	0.08	0.10
6	0.08	0.10	0.11	0.08 (0.27)	0.09 (0.29)	0.09 (0.29)	0.00	0.02	0.06	0.01	0.01	-0.02
7	0.09	0.09	0.11	0.08 (0.27)	0.08 (0.27)	0.09 (0.29)	0.03	0.03	0.07	-0.05	-0.07	-0.07
8	0.07	0.09	0.09	0.07 (0.26)	0.08 (0.27)	0.08 (0.28)	0.00	0.02	0.04	-0.06	-0.08	-0.07
9	0.08	0.07	0.06	0.10 (0.31)	0.10 (0.31)	0.12 (0.32)	-0.09	-0.10	-0.17	-0.10	-0.09	-0.10
10	0.07	0.06	0.04	0.10 (0.30)	0.09 (0.29)	0.09 (0.28)	-0.09	-0.10	-0.17	-0.11	-0.09	-0.07
11	0.07	0.07	0.04	0.09 (0.28)	0.08 (0.28)	0.08 (0.26)	-0.06	-0.07	-0.13	-0.09	-0.09	-0.06
12	0.07	0.06	0.04	0.09 (0.29)	0.09 (0.28)	0.08 (0.27)	-0.10	-0.09	-0.16	-0.05	-0.04	-0.05
English learner student	0.12	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.14 (0.34)	0.01	-0.02	-0.03	-0.01	-0.01	-0.02
Special education student	0.14	0.14	0.14	0.16 (0.37)	0.17 (0.37)	0.17 (0.37)	-0.06	-0.07	-0.07	-0.05	-0.05	-0.04
Fraction of days absent	0.09	0.09	0.07	0.15 (0.20)	0.15 (0.21)	0.13 (0.19)	-0.27	-0.27	-0.32	-0.12	-0.10	-0.10
Number of suspensions	0.14	0.16	0.06	0.18 (0.65)	0.20 (0.74)	0.09 (0.43)	-0.05	-0.05	-0.07	-0.05	-0.05	-0.03
At-risk status	0.42	0.43	0.42	0.42 (0.49)	0.48 (0.50)	0.48 (0.50)	0.00	-0.10	-0.13	0.02	0.01	0.02
Composite of covariates	**3.70**	**3.70**	**3.69**	**3.64 (0.24)**	**3.66 (0.23)**	**3.61 (0.23)**	**0.24**	**0.20**	**0.35**	**0.27**	**0.27**	**0.26**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B10. Exploration of nonresponse bias for social awareness (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with social awareness scale		
	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20
Female	0.50	0.50	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.03	0.04	0.10	0.07	0.06
Grade level												
3	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.34)	0.13 (0.33)	0.10	0.07	0.14	0.11	0.14	0.11
4	0.16	0.16	0.17	0.13 (0.34)	0.13 (0.33)	0.12 (0.33)	0.09	0.08	0.13	0.09	0.11	0.09
5	0.15	0.15	0.16	0.12 (0.32)	0.12 (0.32)	0.11 (0.32)	0.09	0.09	0.13	0.05	0.05	0.08
6	0.08	0.10	0.11	0.08 (0.27)	0.09 (0.29)	0.09 (0.29)	0.00	0.02	0.06	-0.02	-0.02	-0.06
7	0.09	0.09	0.11	0.08 (0.27)	0.08 (0.27)	0.09 (0.29)	0.03	0.03	0.07	-0.08	-0.09	-0.08
8	0.07	0.09	0.09	0.07 (0.26)	0.08 (0.27)	0.08 (0.28)	0.00	0.02	0.04	-0.06	-0.09	-0.08
9	0.08	0.07	0.06	0.10 (0.31)	0.10 (0.31)	0.12 (0.32)	-0.09	-0.10	-0.17	-0.08	-0.08	-0.08
10	0.07	0.06	0.04	0.10 (0.30)	0.09 (0.29)	0.09 (0.28)	-0.09	-0.10	-0.17	-0.08	-0.06	-0.04
11	0.07	0.06	0.04	0.09 (0.28)	0.08 (0.28)	0.08 (0.26)	-0.06	-0.07	-0.13	-0.04	-0.03	-0.02
12	0.07	0.06	0.04	0.09 (0.29)	0.09 (0.28)	0.08 (0.27)	-0.10	-0.08	-0.16	0.00	-0.01	-0.02
English learner student	0.12	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.14 (0.34)	0.01	-0.02	-0.03	0.03	0.00	0.01
Special education student	0.14	0.14	0.14	0.16 (0.37)	0.17 (0.37)	0.17 (0.37)	-0.06	-0.07	-0.07	-0.06	-0.06	-0.07
Fraction of days absent	0.09	0.09	0.07	0.15 (0.20)	0.15 (0.21)	0.13 (0.19)	-0.27	-0.27	-0.32	-0.14	-0.15	-0.13
Number of suspensions	0.14	0.16	0.06	0.18 (0.65)	0.20 (0.74)	0.09 (0.43)	-0.05	-0.05	-0.07	-0.13	-0.12	-0.09
At-risk status	0.42	0.43	0.42	0.42 (0.49)	0.48 (0.50)	0.48 (0.50)	0.00	-0.10	-0.13	-0.04	-0.05	-0.06
Composite of covariates	**3.81**	**3.80**	**3.79**	**3.76 (0.22)**	**3.75 (0.22)**	**3.73 (0.21)**	**0.23**	**0.23**	**0.32**	**0.27**	**0.28**	**0.25**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B11. Exploration of nonresponse bias for rigorous expectations (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with rigorous expectations scale		
	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20
Female	0.50	0.50	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.02	0.03	0.03	0.06	0.04	0.03
Grade level												
3	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.34)	0.13 (0.33)	0.10	0.07	0.13	0.17	0.15	0.12
4	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.33)	0.12 (0.33)	0.09	0.08	0.13	0.16	0.17	0.15
5	0.15	0.15	0.16	0.12 (0.32)	0.12 (0.32)	0.11 (0.32)	0.09	0.08	0.13	0.15	0.15	0.16
6	0.08	0.10	0.11	0.08 (0.27)	0.09 (0.29)	0.09 (0.29)	0.00	0.02	0.06	-0.01	-0.02	-0.05
7	0.09	0.09	0.11	0.08 (0.27)	0.08 (0.27)	0.09 (0.29)	0.02	0.04	0.07	-0.07	-0.09	-0.09
8	0.08	0.09	0.09	0.07 (0.26)	0.08 (0.27)	0.08 (0.28)	0.00	0.02	0.04	-0.09	-0.09	-0.10
9	0.08	0.08	0.06	0.10 (0.31)	0.10 (0.31)	0.12 (0.32)	-0.09	-0.10	-0.17	-0.13	-0.11	-0.13
10	0.07	0.06	0.04	0.10 (0.30)	0.09 (0.29)	0.09 (0.28)	-0.09	-0.10	-0.17	-0.14	-0.13	-0.09
11	0.07	0.07	0.04	0.09 (0.28)	0.08 (0.28)	0.08 (0.26)	-0.06	-0.07	-0.13	-0.12	-0.11	-0.08
12	0.07	0.06	0.04	0.09 (0.29)	0.09 (0.28)	0.08 (0.27)	-0.10	-0.09	-0.16	-0.10	-0.09	-0.09
English learner student	0.12	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.14 (0.34)	0.01	-0.02	-0.03	0.00	-0.01	0.00
Special education student	0.14	0.14	0.14	0.16 (0.37)	0.17 (0.37)	0.17 (0.37)	-0.06	-0.07	-0.07	-0.04	-0.03	-0.04
Fraction of days absent	0.09	0.09	0.07	0.15 (0.20)	0.15 (0.21)	0.13 (0.19)	-0.27	-0.27	-0.32	-0.18	-0.18	-0.14
Number of suspensions	0.14	0.16	0.07	0.18 (0.65)	0.20 (0.74)	0.09 (0.43)	-0.05	-0.05	-0.07	-0.09	-0.10	-0.07
At-risk status	0.42	0.43	0.42	0.42 (0.49)	0.48 (0.50)	0.48 (0.50)	0.00	-0.10	-0.13	0.05	0.03	0.03
Composite of covariates	**4.08**	**4.12**	**4.14**	**4.00 (0.32)**	**4.06 (0.29)**	**4.04 (0.28)**	**0.25**	**0.22**	**0.36**	**0.39**	**0.37**	**0.35**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B12. Exploration of nonresponse bias for student satisfaction (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with student satisfaction scale		
	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20
Female	0.50	0.50	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.02	0.03	0.03	0.01	-0.02	-0.02
Grade level												
3	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.34)	0.13 (0.33)	0.09	0.07	0.13	0.17	0.17	0.19
4	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.33)	0.12 (0.33)	0.09	0.08	0.13	0.14	0.17	0.13
5	0.15	0.15	0.16	0.12 (0.32)	0.12 (0.32)	0.11 (0.32)	0.09	0.08	0.13	0.10	0.10	0.15
6	0.08	0.10	0.11	0.08 (0.27)	0.09 (0.29)	0.09 (0.29)	0.00	0.02	0.06	0.00	-0.01	-0.06
7	0.09	0.09	0.11	0.08 (0.27)	0.08 (0.27)	0.09 (0.29)	0.03	0.04	0.07	-0.07	-0.11	-0.11
8	0.08	0.09	0.09	0.07 (0.26)	0.08 (0.27)	0.08 (0.28)	0.01	0.02	0.04	-0.10	-0.10	-0.12
9	0.08	0.08	0.06	0.10 (0.31)	0.10 (0.31)	0.12 (0.32)	-0.09	-0.10	-0.17	-0.08	-0.08	-0.11
10	0.07	0.06	0.04	0.10 (0.30)	0.09 (0.29)	0.09 (0.28)	-0.09	-0.10	-0.17	-0.11	-0.10	-0.09
11	0.07	0.07	0.04	0.09 (0.28)	0.08 (0.28)	0.08 (0.26)	-0.06	-0.07	-0.13	-0.10	-0.10	-0.09
12	0.07	0.06	0.04	0.09 (0.29)	0.09 (0.28)	0.08 (0.27)	-0.10	-0.09	-0.16	-0.10	-0.10	-0.11
English learner student	0.12	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.14 (0.34)	0.01	-0.01	-0.03	0.10	0.11	0.11
Special education student	0.14	0.14	0.14	0.16 (0.37)	0.17 (0.37)	0.17 (0.37)	-0.06	-0.06	-0.07	-0.03	-0.02	-0.02
Fraction of days absent	0.09	0.09	0.07	0.15 (0.20)	0.15 (0.21)	0.13 (0.19)	-0.27	-0.27	-0.32	-0.16	-0.18	-0.16
Number of suspensions	0.14	0.16	0.07	0.18 (0.65)	0.20 (0.74)	0.09 (0.43)	-0.05	-0.05	-0.07	-0.10	-0.13	-0.09
At-risk status	0.42	0.43	0.42	0.42 (0.49)	0.48 (0.50)	0.48 (0.50)	0.00	-0.10	-0.13	-0.06	-0.06	-0.07
Composite of covariates	**3.75**	**3.75**	**3.76**	**3.69 (0.28)**	**3.69 (0.29)**	**3.67 (0.29)**	**0.22**	**0.19**	**0.32**	**0.36**	**0.38**	**0.38**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B13. Exploration of nonresponse bias for sense of belonging (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with sense of belonging scale		
	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20	2017/18	2018/19	2019/20
Female	0.50	0.50	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.02	0.02	0.03	-0.01	-0.04	-0.05
Grade level												
3	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.34)	0.13 (0.33)	0.09	0.07	0.13	0.16	0.18	0.16
4	0.16	0.15	0.17	0.13 (0.34)	0.13 (0.33)	0.12 (0.33)	0.08	0.08	0.13	0.14	0.16	0.14
5	0.15	0.14	0.16	0.12 (0.32)	0.12 (0.32)	0.11 (0.32)	0.09	0.08	0.13	0.13	0.14	0.17
6	0.08	0.10	0.11	0.08 (0.27)	0.09 (0.29)	0.09 (0.29)	0.01	0.02	0.06	-0.03	-0.04	-0.09
7	0.09	0.09	0.11	0.08 (0.27)	0.08 (0.27)	0.09 (0.29)	0.03	0.04	0.07	-0.08	-0.10	-0.12
8	0.08	0.09	0.09	0.07 (0.26)	0.08 (0.27)	0.08 (0.28)	0.01	0.03	0.04	-0.09	-0.10	-0.10
9	0.08	0.08	0.06	0.10 (0.31)	0.10 (0.31)	0.12 (0.32)	-0.09	-0.10	-0.17	-0.11	-0.12	-0.13
10	0.07	0.06	0.04	0.10 (0.30)	0.09 (0.29)	0.09 (0.28)	-0.09	-0.10	-0.17	-0.12	-0.11	-0.09
11	0.07	0.07	0.04	0.09 (0.28)	0.08 (0.28)	0.08 (0.26)	-0.05	-0.07	-0.13	-0.09	-0.09	-0.08
12	0.07	0.06	0.04	0.09 (0.29)	0.09 (0.28)	0.08 (0.27)	-0.10	-0.08	-0.16	-0.07	-0.07	-0.08
English learner student	0.12	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.14 (0.34)	0.01	-0.01	-0.03	0.05	0.05	0.06
Special education student	0.14	0.14	0.14	0.16 (0.37)	0.17 (0.37)	0.17 (0.37)	-0.06	-0.06	-0.07	-0.01	0.01	0.00
Fraction of days absent	0.09	0.09	0.07	0.15 (0.20)	0.15 (0.21)	0.13 (0.19)	-0.26	-0.27	-0.32	-0.13	-0.15	-0.14
Number of suspensions	0.14	0.16	0.07	0.18 (0.65)	0.20 (0.74)	0.09 (0.43)	-0.05	-0.05	-0.07	-0.05	-0.07	-0.05
At-risk status	0.42	0.43	0.42	0.42 (0.49)	0.48 (0.50)	0.48 (0.50)	0.00	-0.10	-0.13	0.01	0.01	-0.01
Composite of covariates	**3.65**	**3.65**	**3.66**	**3.58 (0.28)**	**3.59 (0.30)**	**3.55 (0.32)**	**0.22**	**0.19**	**0.32**	**0.33**	**0.35**	**0.36**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B14. Exploration of nonresponse bias for changes in perseverance (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with changes in perseverance scale		
	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20
Female	0.50	0.51	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.04	0.05	-0.03	-0.04	-0.09
Grade level												
3	0.21	0.22	0.27	0.15 (0.36)	0.15 (0.36)	0.17 (0.37)	0.16	0.19	0.28	0.01	0.01	0.04
4	0.20	0.21	0.19	0.14 (0.35)	0.14 (0.35)	0.14 (0.34)	0.15	0.21	0.16	0.02	-0.01	-0.03
5	0.14	0.15	0.20	0.11 (0.31)	0.11 (0.31)	0.13 (0.34)	0.09	0.12	0.19	-0.04	-0.04	-0.07
6	0.10	0.14	0.14	0.09 (0.29)	0.11 (0.31)	0.11 (0.31)	0.03	0.09	0.09	-0.04	-0.03	-0.02
7	0.11	0.12	0.06	0.10 (0.30)	0.09 (0.29)	0.10 (0.30)	0.04	0.09	-0.13	0.00	0.03	0.02
8	0.05	0.04	0.04	0.07 (0.25)	0.08 (0.27)	0.08 (0.28)	-0.06	-0.14	-0.17	0.00	0.01	0.01
9	0.06	0.04	0.06	0.11 (0.31)	0.11 (0.31)	0.12 (0.33)	-0.16	-0.22	-0.20	0.00	0.00	0.05
10	0.07	0.04	0.04	0.11 (0.31)	0.10 (0.30)	0.13 (0.33)	-0.14	-0.19	-0.25	0.03	0.04	0.06
11	0.06	0.04	0.00	0.10 (0.31)	0.10 (0.30)	0.02 (0.13)	-0.14	-0.21	-0.13	0.03	0.03	0.01
12	0.00	0.00	0.00	0.01 (0.11)	0.01 (0.10)	0.00 (0.06)	-0.09	-0.09	-0.06	0.00	-0.01	0.00
English learner student	0.13	0.11	0.11	0.13 (0.33)	0.12 (0.33)	0.13 (0.33)	0.00	-0.05	-0.04	0.03	-0.03	-0.02
Special education student	0.12	0.13	0.13	0.16 (0.37)	0.17 (0.37)	0.16 (0.37)	-0.10	-0.10	-0.09	0.02	0.02	0.05
Fraction of days absent	0.07	0.06	0.06	0.12 (0.17)	0.12 (0.18)	0.10 (0.15)	-0.31	-0.35	-0.32	0.02	0.04	0.07
Number of suspensions	0.10	0.11	0.09	0.18 (0.66)	0.20 (0.75)	0.18 (0.67)	-0.12	-0.12	-0.14	0.03	0.01	0.05
At-risk status	0.40	0.39	0.40	0.43 (0.49)	0.47 (0.50)	0.43 (0.50)	-0.05	-0.15	-0.06	0.01	0.01	0.01
Composite of covariates	**-0.05**	**-0.08**	**-0.15**	**-0.03 (0.08)**	**-0.04 (0.10)**	**-0.07 (0.20)**	**-0.20**	**-0.31**	**-0.40**	**0.09**	**0.10**	**0.17**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B15. Exploration of nonresponse bias for changes in self-management (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with changes in self-management scale		
	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20
Female	0.50	0.51	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.04	0.05	-0.02	-0.04	-0.07
Grade level												
3	0.21	0.22	0.28	0.15 (0.36)	0.15 (0.36)	0.17 (0.37)	0.18	0.20	0.29	0.03	0.02	0.03
4	0.20	0.22	0.19	0.14 (0.35)	0.14 (0.35)	0.14 (0.34)	0.16	0.22	0.16	0.01	-0.02	-0.03
5	0.14	0.15	0.20	0.11 (0.31)	0.11 (0.31)	0.13 (0.34)	0.08	0.12	0.19	-0.01	-0.03	-0.05
6	0.10	0.13	0.14	0.09 (0.29)	0.11 (0.31)	0.11 (0.31)	0.02	0.09	0.08	-0.04	-0.04	-0.01
7	0.11	0.12	0.06	0.10 (0.30)	0.09 (0.29)	0.10 (0.30)	0.04	0.08	-0.13	-0.01	0.04	0.01
8	0.05	0.04	0.04	0.07 (0.25)	0.08 (0.27)	0.08 (0.28)	-0.07	-0.14	-0.17	0.00	0.00	0.02
9	0.06	0.04	0.06	0.11 (0.31)	0.11 (0.31)	0.12 (0.33)	-0.16	-0.23	-0.20	-0.01	0.02	0.02
10	0.07	0.04	0.04	0.11 (0.31)	0.10 (0.30)	0.13 (0.33)	-0.14	-0.19	-0.25	0.02	0.02	0.04
11	0.06	0.04	0.00	0.10 (0.31)	0.10 (0.30)	0.02 (0.13)	-0.14	-0.21	-0.13	-0.01	0.01	0.02
12	0.00	0.00	0.00	0.01 (0.11)	0.01 (0.10)	0.00 (0.06)	-0.09	-0.09	-0.06	-0.01	0.00	0.01
English learner student	0.12	0.11	0.11	0.13 (0.33)	0.12 (0.33)	0.13 (0.33)	-0.01	-0.05	-0.05	0.01	0.01	-0.03
Special education student	0.12	0.13	0.13	0.16 (0.37)	0.17 (0.37)	0.16 (0.37)	-0.10	-0.10	-0.09	0.02	0.01	0.03
Fraction of days absent	0.07	0.06	0.06	0.12 (0.17)	0.12 (0.18)	0.10 (0.15)	-0.31	-0.36	-0.32	0.01	0.02	0.07
Number of suspensions	0.10	0.11	0.09	0.18 (0.66)	0.20 (0.75)	0.18 (0.67)	-0.12	-0.12	-0.14	0.03	0.02	0.06
At-risk status	0.40	0.39	0.40	0.43 (0.49)	0.47 (0.50)	0.43 (0.50)	-0.05	-0.15	-0.06	-0.01	-0.01	0.01
Composite of covariates	**0.01**	**-0.05**	**-0.05**	**0.01 (0.06)**	**-0.03 (0.07)**	**0.02 (0.17)**	**-0.06**	**-0.30**	**-0.36**	**0.07**	**0.08**	**0.13**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B16. Exploration of nonresponse bias for changes in self-efficacy (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with changes in self-efficacy scale		
	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20
Female	0.50	0.51	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.04	0.05	-0.03	-0.04	-0.07
Grade level												
3	0.21	0.22	0.28	0.15 (0.36)	0.15 (0.36)	0.17 (0.37)	0.17	0.19	0.29	0.03	0.03	0.09
4	0.20	0.21	0.19	0.14 (0.35)	0.14 (0.35)	0.14 (0.34)	0.16	0.22	0.16	0.02	0.01	-0.04
5	0.14	0.15	0.20	0.11 (0.31)	0.11 (0.31)	0.13 (0.34)	0.08	0.12	0.19	-0.05	-0.07	-0.10
6	0.10	0.14	0.14	0.09 (0.29)	0.11 (0.31)	0.11 (0.31)	0.03	0.09	0.09	-0.05	-0.04	0.00
7	0.11	0.12	0.06	0.10 (0.30)	0.09 (0.29)	0.10 (0.30)	0.04	0.09	-0.13	0.00	0.03	-0.01
8	0.05	0.04	0.04	0.07 (0.25)	0.08 (0.27)	0.08 (0.28)	-0.07	-0.14	-0.17	-0.03	-0.04	0.01
9	0.06	0.04	0.06	0.11 (0.31)	0.11 (0.31)	0.12 (0.33)	-0.16	-0.23	-0.20	0.00	0.01	0.05
10	0.07	0.04	0.04	0.11 (0.31)	0.10 (0.30)	0.13 (0.33)	-0.14	-0.19	-0.25	0.03	0.03	0.05
11	0.06	0.04	0.00	0.10 (0.31)	0.10 (0.30)	0.02 (0.13)	-0.14	-0.21	-0.13	0.04	0.04	0.02
12	0.00	0.00	0.00	0.01 (0.11)	0.01 (0.10)	0.00 (0.06)	-0.09	-0.09	-0.06	0.02	-0.01	0.01
English learner student	0.12	0.11	0.11	0.13 (0.33)	0.12 (0.33)	0.13 (0.33)	-0.01	-0.06	-0.05	0.01	-0.02	-0.03
Special education student	0.12	0.13	0.13	0.16 (0.37)	0.17 (0.37)	0.16 (0.37)	-0.11	-0.10	-0.09	0.02	0.02	0.03
Fraction of days absent	0.07	0.06	0.06	0.12 (0.17)	0.12 (0.18)	0.10 (0.15)	-0.31	-0.36	-0.32	0.04	0.04	0.06
Number of suspensions	0.10	0.11	0.09	0.18 (0.66)	0.20 (0.75)	0.18 (0.67)	-0.12	-0.12	-0.14	0.03	0.01	0.05
At-risk status	0.40	0.39	0.40	0.43 (0.49)	0.47 (0.50)	0.43 (0.50)	-0.05	-0.16	-0.06	0.00	0.01	0.00
Composite of covariates	**-0.05**	**-0.09**	**-0.17**	**-0.03 (0.11)**	**-0.06 (0.12)**	**-0.11 (0.20)**	**-0.20**	**-0.20**	**-0.31**	**0.11**	**0.12**	**0.18**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B17. Exploration of nonresponse bias for changes in social awareness (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with changes in social awareness scale		
	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20
Female	0.50	0.52	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.04	0.05	-0.03	-0.01	-0.06
Grade level												
3	0.21	0.22	0.28	0.15 (0.36)	0.15 (0.36)	0.17 (0.37)	0.17	0.20	0.29	0.01	-0.01	0.03
4	0.20	0.21	0.19	0.14 (0.35)	0.14 (0.35)	0.14 (0.34)	0.16	0.22	0.16	0.00	0.00	-0.08
5	0.14	0.15	0.20	0.11 (0.31)	0.11 (0.31)	0.13 (0.34)	0.08	0.12	0.19	-0.06	-0.07	-0.08
6	0.10	0.14	0.14	0.09 (0.29)	0.11 (0.31)	0.11 (0.31)	0.03	0.09	0.08	-0.03	-0.03	0.00
7	0.11	0.12	0.06	0.10 (0.30)	0.09 (0.29)	0.10 (0.30)	0.04	0.09	-0.13	0.00	0.04	0.04
8	0.05	0.04	0.04	0.07 (0.25)	0.08 (0.27)	0.08 (0.28)	-0.07	-0.14	-0.17	0.00	0.03	0.03
9	0.06	0.04	0.06	0.11 (0.31)	0.11 (0.31)	0.12 (0.33)	-0.16	-0.23	-0.20	0.02	0.03	0.08
10	0.07	0.04	0.04	0.11 (0.31)	0.10 (0.30)	0.13 (0.33)	-0.14	-0.19	-0.25	0.05	0.05	0.09
11	0.06	0.04	0.00	0.10 (0.31)	0.10 (0.30)	0.02 (0.13)	-0.14	-0.21	-0.13	0.04	0.03	-0.01
12	0.00	0.00	0.00	0.01 (0.11)	0.01 (0.10)	0.00 (0.06)	-0.09	-0.09	-0.06	0.01	0.01	-0.01
English learner student	0.12	0.11	0.11	0.13 (0.33)	0.12 (0.33)	0.13 (0.33)	-0.01	-0.05	-0.05	-0.01	0.00	-0.03
Special education student	0.12	0.13	0.13	0.16 (0.37)	0.17 (0.37)	0.16 (0.37)	-0.11	-0.10	-0.10	0.02	0.01	0.01
Fraction of days absent	0.07	0.06	0.06	0.12 (0.17)	0.12 (0.18)	0.10 (0.15)	-0.31	-0.36	-0.32	0.02	0.04	0.06
Number of suspensions	0.10	0.11	0.09	0.18 (0.66)	0.20 (0.75)	0.18 (0.67)	-0.12	-0.12	-0.14	0.04	0.01	0.04
At-risk status	0.40	0.39	0.40	0.43 (0.49)	0.47 (0.50)	0.43 (0.50)	-0.05	-0.15	-0.06	-0.01	-0.02	-0.03
Composite of covariates	**-0.03**	**-0.07**	**-0.13**	**-0.02 (0.09)**	**-0.03 (0.10)**	**-0.05 (0.21)**	**-0.18**	**-0.40**	**-0.39**	**0.11**	**0.12**	**0.19**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B18. Exploration of nonresponse bias for changes in rigorous expectations (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with changes in rigorous expectations scale		
	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20
Female	0.50	0.51	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.04	0.05	-0.04	0.00	-0.06
Grade level												
3	0.21	0.22	0.28	0.15 (0.36)	0.15 (0.36)	0.17 (0.37)	0.16	0.19	0.29	0.04	0.06	0.13
4	0.20	0.21	0.19	0.14 (0.35)	0.14 (0.35)	0.14 (0.34)	0.16	0.21	0.16	0.05	0.03	-0.09
5	0.14	0.15	0.20	0.11 (0.31)	0.11 (0.31)	0.13 (0.34)	0.09	0.12	0.19	-0.12	-0.15	-0.15
6	0.10	0.14	0.14	0.09 (0.29)	0.11 (0.31)	0.11 (0.31)	0.03	0.09	0.09	-0.03	-0.03	0.00
7	0.11	0.12	0.06	0.10 (0.30)	0.09 (0.29)	0.10 (0.30)	0.04	0.09	-0.13	0.02	0.02	0.00
8	0.05	0.04	0.04	0.07 (0.25)	0.08 (0.27)	0.08 (0.28)	-0.06	-0.14	-0.17	0.00	-0.01	0.02
9	0.06	0.04	0.06	0.11 (0.31)	0.11 (0.31)	0.12 (0.33)	-0.16	-0.22	-0.20	0.01	0.01	0.06
10	0.07	0.04	0.04	0.11 (0.31)	0.10 (0.30)	0.13 (0.33)	-0.14	-0.19	-0.25	0.03	0.05	0.07
11	0.06	0.04	0.00	0.10 (0.31)	0.10 (0.30)	0.02 (0.13)	-0.14	-0.20	-0.13	0.02	0.04	0.04
12	0.00	0.00	0.00	0.01 (0.11)	0.01 (0.10)	0.00 (0.06)	-0.09	-0.09	-0.06	0.00	0.00	0.00
English learner student	0.13	0.11	0.11	0.13 (0.33)	0.12 (0.33)	0.13 (0.33)	-0.01	-0.05	-0.05	0.01	-0.01	-0.01
Special education student	0.13	0.13	0.13	0.16 (0.37)	0.17 (0.37)	0.16 (0.37)	-0.10	-0.09	-0.09	0.02	0.01	0.03
Fraction of days absent	0.07	0.06	0.06	0.12 (0.17)	0.12 (0.18)	0.10 (0.15)	-0.30	-0.35	-0.32	0.03	0.05	0.07
Number of suspensions	0.10	0.11	0.09	0.18 (0.66)	0.20 (0.75)	0.18 (0.67)	-0.12	-0.12	-0.14	0.01	0.03	0.03
At-risk status	0.40	0.39	0.40	0.43 (0.49)	0.47 (0.50)	0.43 (0.50)	-0.05	-0.15	-0.06	-0.01	0.00	-0.01
Composite of covariates	**-0.02**	**-0.07**	**-0.14**	**0.00 (0.11)**	**-0.03 (0.14)**	**-0.05 (0.26)**	**-0.15**	**-0.24**	**-0.34**	**0.14**	**0.17**	**0.23**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B19. Exploration of nonresponse bias for changes in student satisfaction (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with changes in student satisfaction scale		
	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20
Female	0.50	0.51	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.04	0.04	-0.03	-0.02	-0.08
Grade level												
3	0.21	0.22	0.27	0.15 (0.36)	0.15 (0.36)	0.17 (0.37)	0.16	0.19	0.28	0.06	0.03	0.13
4	0.20	0.21	0.19	0.14 (0.35)	0.14 (0.35)	0.14 (0.34)	0.15	0.21	0.15	0.03	0.05	-0.08
5	0.14	0.15	0.20	0.11 (0.31)	0.11 (0.31)	0.13 (0.34)	0.08	0.12	0.19	-0.07	-0.10	-0.12
6	0.10	0.14	0.14	0.09 (0.29)	0.11 (0.31)	0.11 (0.31)	0.04	0.09	0.09	-0.07	-0.06	-0.04
7	0.11	0.12	0.06	0.10 (0.30)	0.09 (0.29)	0.10 (0.30)	0.05	0.09	-0.13	0.01	0.03	0.04
8	0.05	0.04	0.04	0.07 (0.25)	0.08 (0.27)	0.08 (0.28)	-0.06	-0.14	-0.17	0.02	0.04	0.05
9	0.06	0.04	0.06	0.11 (0.31)	0.11 (0.31)	0.12 (0.33)	-0.16	-0.22	-0.20	-0.02	-0.01	0.01
10	0.07	0.04	0.04	0.11 (0.31)	0.10 (0.30)	0.13 (0.33)	-0.14	-0.19	-0.25	0.02	0.04	0.06
11	0.06	0.04	0.00	0.10 (0.31)	0.10 (0.30)	0.02 (0.13)	-0.14	-0.20	-0.13	0.01	0.02	0.02
12	0.00	0.00	0.00	0.01 (0.11)	0.01 (0.10)	0.00 (0.06)	-0.09	-0.09	-0.06	0.03	0.01	0.00
English learner student	0.13	0.11	0.11	0.13 (0.33)	0.12 (0.33)	0.13 (0.33)	-0.01	-0.05	-0.05	0.02	-0.01	-0.01
Special education student	0.13	0.13	0.13	0.16 (0.37)	0.17 (0.37)	0.16 (0.37)	-0.09	-0.09	-0.09	0.02	0.02	0.06
Fraction of days absent	0.07	0.06	0.06	0.12 (0.17)	0.12 (0.18)	0.10 (0.15)	-0.30	-0.35	-0.32	0.04	0.04	0.08
Number of suspensions	0.10	0.11	0.09	0.18 (0.66)	0.20 (0.75)	0.18 (0.67)	-0.12	-0.12	-0.14	0.01	0.02	0.03
At-risk status	0.40	0.39	0.40	0.43 (0.49)	0.47 (0.50)	0.43 (0.50)	-0.05	-0.15	-0.06	0.03	0.01	0.04
Composite of covariates	**-0.09**	**-0.13**	**-0.24**	**-0.07 (0.11)**	**-0.10 (0.12)**	**-0.16 (0.22)**	**-0.16**	**-0.25**	**-0.35**	**0.13**	**0.15**	**0.23**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B20. Exploration of nonresponse bias for changes in sense of belonging (student survey)

Covariates and units	Mean for survey sample			Mean for original study sample (with standard deviation)			Difference in standard deviation units			Correlation with changes in sense of belonging scale		
	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20	2017/18 & 2018/19	2018/19 & 2019/20	2017/18 & 2019/20
Female	0.50	0.51	0.51	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.03	0.03	0.04	-0.05	-0.04	-0.10
Grade level												
3	0.21	0.22	0.27	0.15 (0.36)	0.15 (0.36)	0.17 (0.37)	0.16	0.19	0.28	0.03	0.03	0.15
4	0.19	0.21	0.19	0.14 (0.35)	0.14 (0.35)	0.14 (0.34)	0.14	0.20	0.15	0.06	0.07	-0.13
5	0.14	0.15	0.20	0.11 (0.31)	0.11 (0.31)	0.13 (0.34)	0.08	0.12	0.19	-0.14	-0.19	-0.16
6	0.11	0.14	0.14	0.09 (0.29)	0.11 (0.31)	0.11 (0.31)	0.04	0.10	0.10	-0.03	-0.03	0.03
7	0.11	0.12	0.06	0.10 (0.30)	0.09 (0.29)	0.10 (0.30)	0.05	0.09	-0.13	0.02	0.05	0.01
8	0.05	0.04	0.04	0.07 (0.25)	0.08 (0.27)	0.08 (0.28)	-0.05	-0.13	-0.17	-0.02	-0.02	0.01
9	0.06	0.04	0.06	0.11 (0.31)	0.11 (0.31)	0.12 (0.33)	-0.16	-0.22	-0.19	0.01	0.02	0.06
10	0.07	0.04	0.04	0.11 (0.31)	0.10 (0.30)	0.13 (0.33)	-0.14	-0.19	-0.25	0.04	0.05	0.08
11	0.06	0.04	0.00	0.10 (0.31)	0.10 (0.30)	0.02 (0.13)	-0.14	-0.20	-0.13	0.03	0.04	0.01
12	0.00	0.00	0.00	0.01 (0.11)	0.01 (0.10)	0.00 (0.06)	-0.09	-0.09	-0.06	0.02	0.02	0.01
English learner student	0.13	0.11	0.11	0.13 (0.33)	0.12 (0.33)	0.13 (0.33)	-0.01	-0.05	-0.05	0.01	-0.02	-0.02
Special education student	0.13	0.13	0.13	0.16 (0.37)	0.17 (0.37)	0.16 (0.37)	-0.10	-0.09	-0.09	0.02	0.02	0.05
Fraction of days absent	0.07	0.06	0.06	0.12 (0.17)	0.12 (0.18)	0.10 (0.15)	-0.30	-0.35	-0.32	0.04	0.03	0.07
Number of suspensions	0.10	0.11	0.09	0.18 (0.66)	0.20 (0.75)	0.18 (0.67)	-0.12	-0.12	-0.14	0.01	0.02	0.04
At-risk status	0.40	0.40	0.40	0.43 (0.49)	0.47 (0.50)	0.43 (0.50)	-0.05	-0.15	-0.06	0.02	0.00	0.02
Composite of covariates	**-0.06**	**-0.11**	**-0.21**	**-0.03** (0.15)	**-0.08** (0.17)	**-0.13** (0.28)	**-0.16**	**-0.16**	**-0.30**	**0.17**	**0.21**	**0.28**

Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Table B21. Exploration of nonresponse bias for perseverance (teacher survey)

Covariates and units	Mean for survey sample		Mean for original study sample (with standard deviation)		Difference in standard deviation units		Correlation with perseverance scale	
	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19
Female	0.79	0.77	0.77 (0.42)	0.75 (0.43)	0.05	0.05	0.05	0.06
Teacher's years of experience								
0–1 years of experience	0.24	0.21	0.25 (0.43)	0.21 (0.41)	-0.01	-0.02	-0.07	-0.09
2–3 years of experience	0.23	0.21	0.22 (0.42)	0.20 (0.40)	0.01	0.00	-0.03	-0.02
4–5 years of experience	0.11	0.16	0.12 (0.32)	0.15 (0.36)	-0.01	0.01	-0.02	-0.01
6–10 years of experience	0.17	0.18	0.17 (0.38)	0.18 (0.38)	0.01	0.01	0.02	0.01
More than 10 years of experience	0.24	0.25	0.24 (0.43)	0.25 (0.43)	0.00	0.00	0.10	0.09
Teacher's race/ethnicity								
Black	0.50	0.48	0.54 (0.50)	0.50 (0.50)	-0.08	-0.05	0.04	0.00
Hispanic	0.07	0.08	0.07 (0.26)	0.07 (0.26)	0.01	0.01	0.06	0.06
White	0.32	0.34	0.27 (0.45)	0.31 (0.46)	0.10	0.05	-0.10	-0.05
Other	0.04	0.04	0.04 (0.19)	0.04 (0.19)	0.01	0.00	0.03	0.02
Not reported	0.07	0.07	0.08 (0.26)	0.07 (0.25)	-0.03	0.00	0.02	0.01
Type of school								
Elementary	0.72	0.68	0.70 (0.46)	0.67 (0.47)	0.05	0.03	0.25	0.22
Middle	0.29	0.32	0.30 (0.46)	0.32 (0.47)	-0.01	-0.01	-0.06	-0.09
High	0.20	0.22	0.21 (0.41)	0.23 (0.42)	-0.04	-0.03	-0.17	-0.16
Nontraditional	0.00	0.00	0.00 (0.06)	0.00 (0.05)	0.00	0.00	-0.02	0.01
School's STAR rating								
1	0.13	0.14	0.14 (0.34)	0.14 (0.35)	-0.01	-0.01	-0.15	-0.23
2	0.23	0.22	0.23 (0.42)	0.21 (0.41)	-0.02	0.03	-0.12	-0.11
3	0.35	0.19	0.34 (0.47)	0.20 (0.40)	0.02	-0.02	0.07	-0.01
4	0.16	0.30	0.17 (0.37)	0.30 (0.46)	-0.02	-0.01	0.08	0.17
5	0.14	0.16	0.13 (0.33)	0.15 (0.36)	0.02	0.02	0.12	0.14

Table B21. Exploration of nonresponse bias for perseverance (teacher survey) (continued)

Covariates and units	Mean for survey sample		Mean for original study sample (with standard deviation)		Difference in standard deviation units		Correlation with perseverance scale	
	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19
School's ward								
1	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.00	0.02	0.01	-0.03
2	0.06	0.06	0.05 (0.23)	0.06 (0.23)	0.04	0.01	0.05	0.06
3	0.13	0.14	0.12 (0.33)	0.13 (0.34)	0.02	0.02	0.06	0.10
4	0.18	0.16	0.18 (0.38)	0.18 (0.38)	0.01	-0.04	0.05	0.05
5	0.08	0.09	0.09 (0.29)	0.09 (0.29)	-0.03	0.00	-0.06	-0.05
6	0.14	0.15	0.16 (0.37)	0.15 (0.36)	-0.05	-0.01	-0.01	0.06
7	0.12	0.11	0.12 (0.33)	0.12 (0.32)	-0.02	-0.02	-0.01	-0.03
8	0.17	0.16	0.15 (0.36)	0.15 (0.36)	0.04	0.03	-0.08	-0.15
Racial/ethnic composition of the school (fraction of students in the school)								
Black	0.63	0.62	0.65 (0.32)	0.62 (0.32)	-0.04	-0.02	-0.16	-0.19
Hispanic	0.22	0.22	0.21 (0.25)	0.22 (0.24)	0.03	-0.01	0.08	0.07
White	0.11	0.12	0.10 (0.19)	0.12 (0.20)	0.02	0.03	0.12	0.18
Other	0.04	0.04	0.04 (0.05)	0.04 (0.05)	0.04	0.04	0.14	0.20
School-level averages across students								
ELA achievement test score	-0.09	-0.07	-0.10 (0.56)	-0.07 (0.57)	0.01	0.01	0.20	0.26
Math achievement test score	-0.07	-0.05	-0.08 (0.54)	-0.06 (0.54)	0.02	0.01	0.20	0.26
Fraction of days absent	0.89	0.89	0.89 (0.11)	0.89 (0.12)	0.01	0.01	0.22	0.25
Number of suspensions	0.12	0.13	0.13 (0.15)	0.13 (0.18)	-0.03	-0.02	-0.24	-0.26
Composite of covariates	**3.20**	**3.21**	**3.19 (0.29)**	**3.21 (0.31)**	**0.01**	**0.00**	**0.37**	**0.40**

ELA is English language arts. STAR is School Transparency and Reporting Framework.
Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table B22. Exploration of nonresponse bias for rigorous expectations (teacher survey)

	Mean for survey sample		Mean for original study sample (with standard deviation)		Difference in standard deviation units		Correlation with rigorous expectations scale	
Covariates and units	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19
Female	0.79	0.77	0.77 (0.42)	0.75 (0.43)	0.05	0.05	0.10	0.10
Teacher's years of experience								
0–1 years of experience	0.24	0.20	0.25 (0.43)	0.21 (0.41)	-0.01	-0.02	-0.11	-0.09
2–3 years of experience	0.23	0.20	0.22 (0.42)	0.20 (0.40)	0.02	0.00	-0.01	-0.02
4–5 years of experience	0.11	0.16	0.12 (0.32)	0.15 (0.36)	-0.01	0.01	0.02	0.01
6–10 years of experience	0.17	0.18	0.17 (0.38)	0.18 (0.38)	0.00	0.01	0.02	0.01
More than 10 years of experience	0.24	0.25	0.24 (0.43)	0.25 (0.43)	0.00	0.00	0.09	0.08
Teacher's race/ethnicity								
Black	0.50	0.48	0.54 (0.50)	0.50 (0.50)	-0.08	-0.05	0.11	0.10
Hispanic	0.07	0.08	0.07 (0.26)	0.07 (0.26)	0.01	0.01	-0.01	0.03
White	0.32	0.34	0.27 (0.45)	0.31 (0.46)	0.10	0.05	-0.11	-0.13
Other	0.04	0.04	0.04 (0.19)	0.04 (0.19)	0.01	0.00	-0.04	-0.02
Not reported	0.07	0.07	0.08 (0.26)	0.07 (0.25)	-0.03	0.00	0.03	0.01
Type of school								
Elementary	0.72	0.68	0.70 (0.46)	0.67 (0.47)	0.05	0.03	0.10	0.12
Middle	0.29	0.32	0.30 (0.46)	0.32 (0.47)	-0.01	-0.01	-0.02	0.01
High	0.20	0.22	0.21 (0.41)	0.23 (0.42)	-0.04	-0.02	-0.09	-0.13
Nontraditional	0.00	0.00	0.00 (0.06)	0.00 (0.05)	0.00	0.00	-0.02	0.00
School's STAR rating								
1	0.13	0.14	0.14 (0.34)	0.14 (0.35)	-0.02	-0.01	-0.03	-0.08
2	0.23	0.22	0.23 (0.42)	0.21 (0.41)	-0.02	0.03	-0.01	0.00
3	0.35	0.19	0.34 (0.47)	0.20 (0.40)	0.02	-0.03	0.03	0.02
4	0.16	0.30	0.17 (0.37)	0.30 (0.46)	-0.02	-0.02	-0.03	0.05
5	0.14	0.16	0.13 (0.33)	0.15 (0.36)	0.03	0.03	0.03	-0.01

Table B22. Exploration of nonresponse bias for rigorous expectations (teacher survey) (continued)

	Mean for survey sample		Mean for original study sample (with standard deviation)		Difference in standard deviation units		Correlation with rigorous expectations scale	
Covariates and units	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19
School's ward								
1	0.12	0.13	0.12 (0.32)	0.12 (0.33)	0.00	0.02	0.00	-0.04
2	0.06	0.06	0.05 (0.23)	0.06 (0.23)	0.04	0.01	-0.01	-0.03
3	0.13	0.14	0.12 (0.33)	0.13 (0.34)	0.02	0.02	-0.03	0.00
4	0.18	0.16	0.18 (0.38)	0.18 (0.38)	0.01	-0.04	0.02	0.04
5	0.08	0.09	0.09 (0.29)	0.09 (0.29)	-0.03	0.00	-0.02	-0.01
6	0.14	0.15	0.16 (0.37)	0.15 (0.36)	-0.05	-0.01	-0.01	0.04
7	0.12	0.11	0.12 (0.33)	0.12 (0.32)	-0.02	-0.02	0.05	0.04
8	0.17	0.16	0.15 (0.36)	0.15 (0.36)	0.04	0.03	0.00	-0.05
Racial/ethnic composition of the school (fraction of students in the school)								
Black	0.63	0.62	0.65 (0.32)	0.62 (0.32)	-0.04	-0.02	0.01	-0.01
Hispanic	0.22	0.22	0.21 (0.25)	0.22 (0.24)	0.03	-0.01	0.00	0.01
White	0.11	0.12	0.10 (0.19)	0.12 (0.20)	0.02	0.03	-0.01	0.00
Other	0.04	0.04	0.04 (0.05)	0.04 (0.05)	0.04	0.04	-0.01	0.00
School-level averages across students								
ELA achievement test score	-0.09	-0.07	-0.10 (0.56)	-0.07 (0.57)	0.01	0.01	0.02	0.04
Math achievement test score	-0.07	-0.05	-0.08 (0.54)	-0.06 (0.54)	0.02	0.01	0.02	0.02
Fraction of days absent	0.89	0.89	0.89 (0.11)	0.89 (0.12)	0.01	0.01	0.09	0.11
Number of suspensions	0.12	0.13	0.13 (0.15)	0.13 (0.18)	-0.04	-0.02	-0.07	-0.07
Composite of covariates	**4.57**	**4.56**	**4.57 (0.10)**	**4.56 (0.11)**	**-0.01**	**-0.01**	**0.23**	**0.25**

ELA is English language arts. STAR is School Transparency and Reporting Framework.
Note: This table explores nonresponse bias through the mean, standard deviation, and correlation with the scale across covariates.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table B23. Difference between survey sample and original sample for the composite of covariates with and without nonresponse weights

Measure	Difference in composite of covariates between survey sample and original sample in standard deviation units											
	2017/18		2018/19		2019/20		2017/18 & 2018/19		2018/19 & 2019/20		2017/18 & 2019/20	
	Without weights	With weights	Without weights	With weights	Without weights	With weights	Without weights	With weights	Without weights	With weights	Without weights	With weights
(a) Student survey												
Perseverance	0.19	-0.03	0.13	-0.02	0.22	0.00	-0.20	0.07	-0.31	0.10	-0.40	0.04
Self-management	0.18	-0.03	0.19	-0.03	0.23	-0.04	-0.06	0.05	-0.30	0.06	-0.36	0.04
Self-efficacy	0.24	-0.02	0.20	-0.02	0.35	-0.03	-0.20	0.09	-0.20	0.09	-0.31	0.03
Social awareness	0.23	-0.03	0.23	-0.03	0.32	-0.05	-0.18	0.07	-0.40	0.06	-0.39	0.05
Rigorous expectations	0.25	-0.03	0.22	-0.03	0.36	-0.04	-0.15	0.05	-0.24	0.09	-0.34	0.00
Student satisfaction	0.22	-0.03	0.19	-0.03	0.32	-0.07	-0.16	0.07	-0.25	0.07	-0.35	0.05
Sense of belonging	0.22	-0.03	0.19	-0.03	0.32	-0.05	-0.16	0.05	-0.16	0.07	-0.30	0.06
(b) Teacher survey												
Perseverance	0.01	0.00	0.00	0.01	na	na	na	na	na	na	na	na
Rigorous expectations	-0.01	0.00	-0.01	0.00	na	na	na	na	na	na	na	na

na is not applicable.

Note: This table displays the difference in the composite of covariates between the survey sample and original sample, with and without nonresponse weights, in standard deviation units.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Research question 1. How do average SEL competencies and school experiences differ across grade levels and change for individual students between years? *Do student and teacher reports of school experiences and SEL competencies change in similar ways across grade levels? To what extent do the average differences in students' SEL competencies and school experiences across grades differ by the type of students (such as students classified by gender, race/ethnicity, and academic achievement)? How are individual students' reports of SEL competencies and school experiences associated between years, and how does that association compare to that of other variables (such as achievement test scores, absences, and suspensions)?*

Using nonresponse weights, the study team calculated the average value of students' self-reported SEL competencies and school experiences for each grade and age for the full sample of students. The analyses were conducted separately by grade and age to account for the possibility that students might repeat grades so that not all students in a given grade will be the same age. Using nonresponse weights, the study team calculated the average value of teachers' reports of SEL competencies and school experiences by grade level of the teachers' school (elementary, middle, or high school). To provide a comparison to the teacher averages, the study team also calculated averages of the students' reports of SEL competencies and school experiences by school level. To assess whether the measures varied across grades and school levels, the study team conducted *F*-tests of the null hypothesis that the measures were equal across grades and school levels. The resulting *p*-value provided evidence about whether the measures differed significantly across grades.

The student analyses were also conducted separately by student characteristics, including gender, cohort, race/ethnicity, geographic ward, special education status, English learner status, at-risk status, performance on achievement tests, suspension status, and chronic absenteeism status. The subgroups were defined based on the time of the reporting for a particular year.

Other analyses examined year-to-year correlations between SEL competencies and school experiences, as well as academic outcomes, shedding light on the degree to which individual students' reports are consistent year-to-year. To do so, the study team calculated pairwise correlations between two measurements of the same SEL competency or school experiences in consecutive years, using nonresponse weights. To provide information on the direction of the relationship, the correlations were not constrained to be positive. If a correlation is closer to 0, it suggests that students' SEL measures change from year to year because the measure in one year is not highly correlated with the measure in another year. These statistics were calculated separately by student characteristics, including gender, race/ethnicity, geographic ward, special education status, English learner status, at-risk status, performance on achievement tests, suspension status, and chronic absenteeism status.

Research question 2. To what extent do year-to-year changes in individual students' SEL competencies and school experiences differ across schools?

To examine the variation in measures within versus across schools, the study team calculated the intraclass correlation coefficient (ICC), a statistic that represents the extent to which a variable fluctuates within groups versus across groups. The study team calculated the ICC for SEL competencies and school experiences, as well as changes in SEL competencies and school experiences across years. The results provided a way to decompose the overall variance into within- and between-school components. In particular, the ICC is the fraction of a variable's total variance that arises across schools, so the units are comparable across different variables.[22] To provide a benchmark, ICCs were calculated for other academic measures, including math and ELA achievement, in-seat attendance, and number of suspensions. For the analyses of changes in measures, the school assignment was

[22] In particular, the ICC is invariant to linear changes in scale. For example, the ICC would not change if a variable were multiplied by a constant. However, the ICC is not invariant to nonlinear changes in scale. For example, the ICC may change if the natural log of the variable were used instead of the original variable.

based on the second year of the measurement, because that is the school that (typically) students will have attended the longest between when the two measures were collected. When calculating the ICCs, nonresponse weights were used.

The ICCs for changes in measures were translated into a measure that represents the expected difference in percentile points between students who attend a school with a high average change for each measure compared to a school with an average change for each measure. This translation was designed to provide a more intuitive metric for interpreting the results than the ICC. A high-change school was defined as being one standard deviation above average in terms of change within DCPS. Although this metric was presented as a comparison of student growth between two types of schools, it was based on the ICC as calculated using the full sample of schools.

To make the translation, the study team applied the following formula:

$$Difference\ in\ percentiles = 100 \times \Phi\left(\sqrt{ICC} \times \sqrt{2[1 - corr(M_{t+1}, M_t)]}\right) - 50,$$

where Φ is the standard normal cumulative distribution function and $corr(M_{t+1}, M_t)$ is the correlation between the measure in successive years.

The formula is based on translating the ICC into standard deviation units. Let σ_B^2 be the variance of change in the measure between schools, σ_T^2 be the total variance in change in the measure, and σ_M^2 be the total variance of the levels of the measure.

By construction, $ICC = \sigma_B^2/\sigma_T^2$, which implies that $\sigma_B = \sqrt{ICC} \times \frac{\sigma_T}{\sigma_M} \times \sigma_M$. In other words, one standard deviation of change in the measure between schools is equivalent to $\sqrt{ICC} \times \frac{\sigma_T}{\sigma_M}$ standard deviations of the measure.

Define a high-change school as one with change that is one school-level standard deviation above the mean and a typical school as one with change at the mean. For two students who initially start off with the same level of the measure, the student who attends the high-change school is expected to have a measure that is $\sqrt{ICC} \times \frac{\sigma_T}{\sigma_M}$ standard deviations higher after one year than the student who attends the typical school. Suppose that the student in the typical school has an outcome (M_{t+1}) that is equal to the mean of the outcome after one year (μ). Assuming that M_{t+1} follows a normal distribution, $M_{t+1} \sim N(\mu, \sigma_M^2)$, then the student's outcome in the typical school is at the 50th percentile. The outcome of the student in the high-change school is then expected to be $\mu + \sqrt{ICC} \times \frac{\sigma_T}{\sigma_M}\sigma_M$. The percentile for the student in the high-change school is therefore given by $100 \times \Phi\left(\sqrt{ICC} \times \frac{\sigma_T}{\sigma_M}\right)$, where $\Phi()$ is the standard normal cumulative distribution function. Because $\frac{\sigma_T}{\sigma_M} = \sqrt{2[1 - corr(M_{t+1}, M_t)]}$, the difference in percentiles between the student in the high-change school and typical school can be written as:

$$Difference\ in\ percentiles = 100 \times \Phi\left(\sqrt{ICC} \times \sqrt{2[1 - corr(M_{t+1}, M_t)]}\right) - 50.$$

Research question 3. How do measures of SEL competencies and school experiences relate to future outcomes, and how do they complement other available data for predicting future outcomes? *To what extent do individual SEL competencies and school experiences relate to student outcomes measured one and two years later (such as achievement test scores, absences, suspensions, and whether a student feels loved, challenged, and prepared)? When other data are available—such as demographic information, achievement test scores, absences, and suspensions—to predict students' future outcomes, to what extent does adding measures of SEL competencies and school experiences improve the predictive power and accuracy of those predictions? Which types of data and statistical models could best help DCPS classify whether students are at risk of having negative outcomes?*

Predictive power of individual predictors

To gauge the predictive power of each predictor, the analyses included models that examine the relationship between each key outcome and each predictor variable. The predictors included SEL competencies, school experiences, and academic measures (achievement test scores, in-seat attendance, and number of suspensions). The study team summarized predictive power using two complementary statistics: pairwise correlations and classification accuracy.

1. **Pairwise correlations.** The study team used pairwise correlations to summarize the strength of a predictor and an outcome. All predictor variables were continuous. The correlations were calculated using nonresponse weights.
2. **Classification accuracy of individual predictors.** To provide a measure of how well individual predictors could identify students at risk of poor outcomes, the study team estimated the predictor's classification accuracy. For binary outcomes, predictor variables can be used to classify whether a student is likely to achieve an outcome or not—for example, whether they are likely to graduate from high school or not. A good predictor would have a higher classification accuracy in that it correctly classified a higher fraction of outcomes.

For outcome variables with little variation, total classification accuracy can be misleading. For example, consider an outcome that 95 percent of students achieve with success. Without additional information, classifying all students as successful would result in a relatively high classification accuracy of 95 percent. The baseline accuracy (accuracy of the null model) is the percentage of students who would be correctly classified if all students were classified with the most prevalent value for each outcome (in this example, 95 percent). Therefore, a predictor that classified 95 percent of students as successful would not add much value. For this reason, the study team examined improvements in classification accuracy relative to baseline accuracy. For example, if 95 percent of students achieved a positive outcome and the predictor yielded a 97 percent classification accuracy, the improvement in classification accuracy would be 2 percentage points.

The study team calculated classification accuracy in five steps:

1. The study team converted any continuous outcomes into meaningful binary variables. For example, in-seat attendance was converted into whether a student was chronically absent or not.
2. The study team estimated probit models of the outcome as a function of the predictor. This model provided an estimate of the probability that each student would have a given outcome (for example, the probability that a student would be chronically absent). The model was estimated using nonresponse weights.
3. Based on the estimated probabilities, the study team determined a cutoff in terms of the probabilities that would maximize the classification accuracy, such that students with an estimated probability above or below the cutoff would be classified as likely to have that outcome or not. Using nonresponse weights, the classification accuracy was calculated for potential cutoffs in increments of 0.01. The study team selected the cutoff that yielded the highest classification accuracy out of those that were tested. The classification accuracy associated with this cutoff was reported as the total classification accuracy (ACC).
4. Using nonresponse weights, the study team calculated the percentage of students who had a value of 1 for the outcome ($\hat{P}$).
5. The improvement in accuracy ($IACC$) was calculated by subtracting from the classification accuracy the baseline accuracy of classifying all students the same way given information on $\hat{P}$: $IACC = ACC - \max(\hat{P}, 1 - \hat{P})$.

Predictive power of groups of variables

The study also considered the predictive power of groups of variables by estimating multivariate models that examine the relationship between students' outcomes and groups of the predictor variables. The analyses examined the predictive power of all SEL competencies and school experiences combined, all academic measures combined, and all demographic characteristics combined. The analyses also examined the predictive power of combinations of the groups of variables, which provides information on the incremental predictive power—the extent to which one group of variables adds to the predictive power of another group of variables. To assess the predictive power of groups of variables, the study team adopted three approaches:

1. **Multivariate correlation**. A multivariate correlation is a measure of strength between a single variable of interest and a group of other variables. It ranges from 0 to 1. The higher the multivariate correlation, the more related the variable is to the group of other variables. The study team calculated the multivariate correlation as the square root of the adjusted *R*-squared from an ordinary least squares regression of each outcome on a group of predictors. The regressions accounted for nonresponse weights. The incremental predictive power is the difference in multivariate correlations between different groups of predictors. For example, the incremental predictive power of the SEL competencies and school experiences is the difference between the multivariate correlation based on a regression that uses all the predictors and a multivariate correlation based on a regression that uses all the predictors except the SEL competencies and school experiences. The predictor variables included a combination of continuous and categorical variables. For categorical variables, the models included indicators for each category, excluding a reference category.
2. **Classification accuracy of groups of variables based on probit models.** Using the methodology described in the analyses of individual predictors, the study team estimated the percentage of students that each group of predictors could correctly classify. Each predictor was included separately in the model. In other words, the models did not include interaction terms between the predictors. To assess the incremental predictive power, the accuracies were compared across groups of predictors. For example, the incremental predictive power of the SEL competencies and school experiences is the difference between the percentage that are correctly classified from the model that uses all the predictors and the percentage that are correctly classified from the model that uses only the predictors based on variables other than the SEL competencies and school experiences. This metric provides a practical sense of how much the inclusion of SEL competencies and school experiences could boost the ability to predict student outcomes.
3. **Classification accuracy of groups of variables based on random forest models.** Random forests are a machine learning method commonly used for prediction problems. The study team implemented random forest models using the same groups of variables as in the probit approach, including SEL competencies, school experiences, academic measures, and basic demographics. The models were estimated with the *ranger* package for the statistical programming language R (Wright & Ziegler, 2017). The study team used standard procedures for selecting the options for the analysis. Specifically, the study team constructed 500 decision trees for each model. For each decision tree, the study team sampled data with replacement and evaluated the square root of the number of available predictors at each split. The modeling approach was designed to align with the probit methodology (for example, over-sampling observations with larger nonresponse weights) in order to compare the incremental predictive power afforded by different sets of predictors.[23]

[23] The random forest predictions are based on out-of-bag samples—the predictions of the decision trees in which a given case was *not* included in the bootstrapped sample. As a sensitivity, the team also estimated the random forest models using 10-fold cross-validation. These cross-validated models produced results which closely aligned with the relative accuracy rates of the benchmark approach, always falling within ± 3 percentage points.

Linking "credits behind in grade 9" to high school graduation

The study team used a similar type of analysis to identify the extent to which grade 8 students are likely to fall behind academically, focusing on the middle and high school sample. This analysis provides context for a key subgroup analysis that examined how grade 8 SEL competencies and school experiences relate to credits earned in high school. To understand which SEL competencies and school experiences might be most important for transitioning from grade 8 to high school, the study team estimated pairwise correlations between the SEL competencies and school experiences of grade 8 students in 2017/18 and the extent to which students were behind in terms of credits earned in 2018/19. DCPS currently predicts whether students are falling behind using a measure based on the number of credits by which a student is behind for his or her grade level (credits behind).

Ideally, the study would have examined the probability of graduating from high school given data from grade 8, but the study only had access to three years of data, so it is not possible to calculate this probability directly. Instead, using data on credits behind from the 2017/18 and 2018/19 school years and graduation data from the 2018/19 school year, the study examined the extent to which credits behind in grade 9 reflected the likelihood of graduating by linking data from several cohorts. This analysis provides a sense of whether the grade-8 SEL competencies and school experiences are related to high school graduation through their link to the number of credits behind in grade 9. For example, if grade-8 SEL competencies relate to grade-9 credits behind and grade-9 credits behind relate to high school graduation, then grade-8 SEL competencies could be an indicator for likely high school graduation. One limitation of this approach is that it assumes that the models are the same across subsequent cohorts of students. The analyses proceeded in two steps:

Step 1: Establishing the relationship between credits behind in grade 9 and the probability of graduating high school. Using ordinary least squares regressions, the study team estimated the relationship between credits behind in each grade as a function of credits behind in the previous grade. In addition, using a probit model, the study team estimated the probability of graduating given the number of credits behind in grade 11. This information was combined to form a probability of graduating given the grade-9 credits behind. Formally, the following equations were estimated using data from several separate cohorts of students:

(Equation B1)

$$
\begin{aligned}
D^{*}_{i12} &= \alpha_{11} + \rho_{11} M_{i11} + \nu_{i11} \\
M_{i11} &= \alpha_{10} + \rho_{10} M_{i10} + \nu_{i10} \\
M_{i10} &= \alpha_{09} + \rho_{09} M_{i09} + \nu_{i09},
\end{aligned}
$$

where M_{it} is the number of credits behind at time t, D^{*}_{i12} is the propensity to graduate from high school, and ν_{i11}, ν_{i10}, and ν_{i09} are error terms assumed to be independent from each other. Each equation was estimated separately for a single cohort using data from the 2017/18 and 2018/19 school years. For example, ρ_{09} was estimated using the cohort of students who were in grade 9 during the 2017/18 school year. Estimates of α_t and ρ_t were recursively substituted to relate D^{*}_{i12} to M_{i9}.

Step 2: Estimating the extent to which credits behind in grade 9 can accurately classify high school graduation. The study team conducted a simulation to determine the predictive accuracy of grade-9 credits behind in classifying high school graduation. Because the sample did not include students in grade 9 who could have graduated in the timeframe of the study, it was not possible to directly calculate the predictive accuracy of grade-9 credits behind. Instead, the study team conducted a simulation using the following steps:

1. The data were subset to students in grade 9 in 2017/18.
2. Using the estimated residual variance of ν_{i09} ($\hat{\sigma}^2_{\nu 09}$), a simulated error term was drawn for each student from the normal distribution: $N(0, \hat{\sigma}^2_{\nu 09})$.
3. Based on estimates of α_{09} and ρ_{09} and the simulated error, a simulated value of credits behind in grade 10 was formed for each student ($\widehat{M}_{i10}$).

4. Using $\widehat{M}_{i10}$, step (1) was repeated step to form simulated values of grade-11 credits earned and high school graduation.
5. Then, treating the simulated data as real, the estimated values of equation B1 were used to classify students as likely to graduate or not based on their grade-9 credits behind. This step provided an estimate of the accuracy of grade-9 credits behind in classifying high school graduation.
6. Steps (1) through (5) were repeated 500 times to form 500 estimates of the classification accuracy. The average across those estimates was used to form an overall classification accuracy.

Research question 4. How do measures of perseverance and rigorous expectations align across students, parents, and teachers? *Across schools, to what extent do survey reports on these measures from students, parents, and teachers align? Is alignment associated with characteristics of schools (such as the demographic characteristics of their student population and the schools' accountability ratings) and response rates on the survey?*

Because of data limitations, the analyses focused on estimating correlations among school-level averages of SEL competencies and school experiences among respondent types. Ideally, the analysis would have examined the alignment between students, parents, and teachers at the student level, but this analysis is not possible for two reasons: teacher reports about the SEL competencies of groups of students (not individual students) and parent responses are not linked to their child.[24] Instead, the analyses examined alignment at the school level by forming averages of the reports from each type of respondent for each school. The study team then calculated pairwise correlations among the averages.

To examine overall alignment, the study team estimated the average student, parent, and teacher reports for each school. The correlations among the measures suggest whether the responses by respondent type tend to move together, but these correlations do not inform whether the overall responses align in terms of their levels. To do so, the analyses explored two types of measures:

1. **The difference between each pair of respondent types.** These values could be negative or positive, so they indicate which respondent type had the higher value.
2. **The absolute difference between pairs of respondent types.** The absolute differences are constrained to be positive so they indicate the magnitude of the differences but do not indicate which respondent types had the higher value.

The analyses examined the extent to which these measures of alignment vary across schools and whether alignment relates to other characteristics of schools, including student composition (percentage of students by race/ethnicity, English learner status, and special education status), STAR rating, geographic ward, and response rates on the surveys. For continuous school characteristics (for example, percentage of students in various demographic groups), the study team calculated correlations between the measures of misalignment and the school characteristic. For categorical school characteristics (for example, geographic ward), the study team calculated the levels of misalignment separately for each category.

The study team did not adjust for nonresponse in these analyses, because it was not possible to explore nonresponse bias for parents. Because DCPS does not have information on the number of parents eligible for the survey, it is not possible to calculate response rates on the parent survey. For consistency, the study team did not

[24] The items in the teacher survey ask teachers to rate the competencies of the students that they teach. Therefore, teachers in different types of schools will rate different groups of students. For example, elementary school teachers might rate a group of students they teach for all school subjects, whereas high school teachers might rate the competencies of students within a particular subject.

use student or teacher nonresponse weights for results reported in the main text. As a sensitivity check, the study team conducted analyses using student and teacher nonresponse weights to confirm the findings generalize to the population of respondents at each school.

Benchmarks

To describe the strength of the correlations in this context, this study adopted benchmarks based on past evidence on the correlations between cognitive tests (IQ tests and achievement tests) and other future academic outcomes. Historically, cognitive tests have been most widely used in educational assessment (Heckman & Kautz, 2014), so their predictive power serves as a natural comparison for other types of assessments that are starting to be used more frequently, including measures of SEL competencies and school experiences.

This study described ranges of correlations as follows: 0.0 to 0.09 is low, 0.10 to 0.19 is moderate, 0.20 to 0.29 is substantive, and 0.30 and above is high. The study team viewed correlations above 0.30 as high, because this value represents the upper end of the estimated correlations between cognitive tests and other future academic outcomes from prior research (table B24). Given that many estimates of the correlations between cognitive tests and other academic outcomes fell in the range of 0.20 to 0.29, the study viewed this range as substantive. Correlations in the range of 0.10 to 0.19 were classified as moderate because they were approximately half the magnitude of those classified as high and, thus, may still be meaningful.

Table B24. Correlations between cognitive tests and other future academic outcomes in existing literature

Type of test	Outcome	Correlation	Source
Achievement test	High school graduation	0.33	Kautz & Zanoni (2020)
Achievement test	College grades	0.15-0.28[a]	Noftle & Robins (2007)
Achievement test	Graduated college	0.27	Galla et al. (2019)
Achievement test	Completed a bachelor's degree	0.43	Heckman & Kautz (2012)
IQ test	Completed a bachelor's degree	0.35	Heckman & Kautz (2012)
IQ test	Academic performance	0.23	Poropat (2009)
IQ test	Years of educational attainment	0.07-0.20[a]	Almlund et al. (2011)

[a] Based on standardized regression coefficients, which have a similar interpretation to a correlation.

To describe the magnitude of differences in SEL competencies and school experiences between groups of students, this study adopted benchmarks based on past evidence on the effect of school-based SEL programs. This evidence provides a natural benchmark in this context because it suggests the extent to which SEL programming could reduce differences between groups of students. Therefore, differences that exceed the effects of school-based SEL programs could be viewed as large because the differences might be challenging to eliminate through SEL programming.

To set the benchmarks, the study team drew on a meta-analysis of school-based SEL programs. The meta-analysis estimated effects in standard deviation units of 0.57 on SEL skills, 0.23 on attitudes, 0.24 on positive social behavior, 0.22 on conduct problems, and 0.24 on emotional distress (Durlak et al., 2011). Because these categories of outcomes span the SEL competencies and school experience variables in this study, the average effect—0.30 standard deviations—provided a relevant summary of the overall effect. The study team, therefore, described differences of 0.30 standard deviations and above as large. For consistency with the benchmarks for correlations, the study described smaller differences between groups as follows: 0.0 to 0.09 standard deviations is small, 0.10 to 0.19 standard deviations is moderate, and 0.20 to 0.29 is substantive.

Limitations

Four main limitations should be kept in mind. First, as discussed in the main text, the study examined correlational relationships between variables, not causal ones. Therefore, the findings do not indicate that improvements in SEL competencies or school experiences would necessarily improve students' outcomes.

Second, the study focused on two years of data on each students' SEL competencies and school experiences, which affects the interpretation of the findings. The analyses of how SEL competencies and school experiences evolve with grade relied on comparing different students in different cohorts and grades rather than examining changes in the same student over time. This approach requires that SEL competencies do not evolve in different ways across cohorts of students. Supporting this requirement, the trends were similar across the two cohorts in the analysis sample. As discussed in the main text, the study cannot rule out that the U-shaped pattern emerged because the composition of students changed across grades. Although future research is needed to determine the source or sources of this pattern, the results are still useful for DCPS, because the estimates apply to the students who are enrolled and who DCPS can still reach through programming or policies. For example, DCPS can use these estimates to help inform decisions about when to target supports for enrolled students.

Third, data limitations reduced the generalizability of the study in several ways. The study focused on one school district during a relatively short timeframe and on one of many possible surveys that school districts could use, potentially limiting the generalizability. Because the surveys were administered recently, the study could not examine longer-term trends and outcomes, and the measure of alignment across respondent types could reflect differences in survey response rates (which cannot be adjusted for, given a lack of data on parents).

Fourth, the COVID-19 pandemic affected both the administrative and survey data from the 2019/20 school year, which limited its use in this study. The administration of the survey, which is typically administered online to students in school in early March, was only available online for students to complete at home once students could no longer attend school in-person—resulting in a lower response rate for students. For all respondents, the survey was not likely a top priority given the pandemic's disruption of normal life, and those who did respond might have responded differently as a result. For these reasons, the main analyses did not include the 2019/20 survey data. In addition, certain annual standardized tests were canceled in spring 2020, including the PARCC exams and spring administrations of the SAT and PSAT.

References

Almlund, M., Duckworth, A., Heckman, J. J., & Kautz, T. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 1–181). Elsevier.

District of Columbia Office of the State Superintendent of Education. (2020, August 4). *2020 DC school report card & STAR Framework technical guide.* https://osse.dc.gov/sites/default/files/dc/sites/osse/publication/attachments/2020%20Report%20Card%20and%20STAR%20Framework%20Technical%20Guide.pdf

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432.

Galla, B. M., Shulman, E. P., Plummer, B. D., Gardner, M., Hutt, S. J., Goyer, J.P., et al. (2019). Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *American Educational Research Journal, 56*(6), 2077–2115.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*(4), 451–464.

Heckman, J. J., & Kautz, T. (2014). Achievement tests and the role of character in American life. 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. 3–56). University of Chicago Press.

Kautz, T., & Zanoni, W. (2020). *Measuring and fostering non-cognitive skills in adolescence: Evidence from Chicago Public Schools and the OneGoal Program.* Unpublished manuscript.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93*(1), 116–130.

Poropat, A. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338.

Wright, M., & Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1), 1–17. https://www.jstatsoft.org/article/view/v077i01

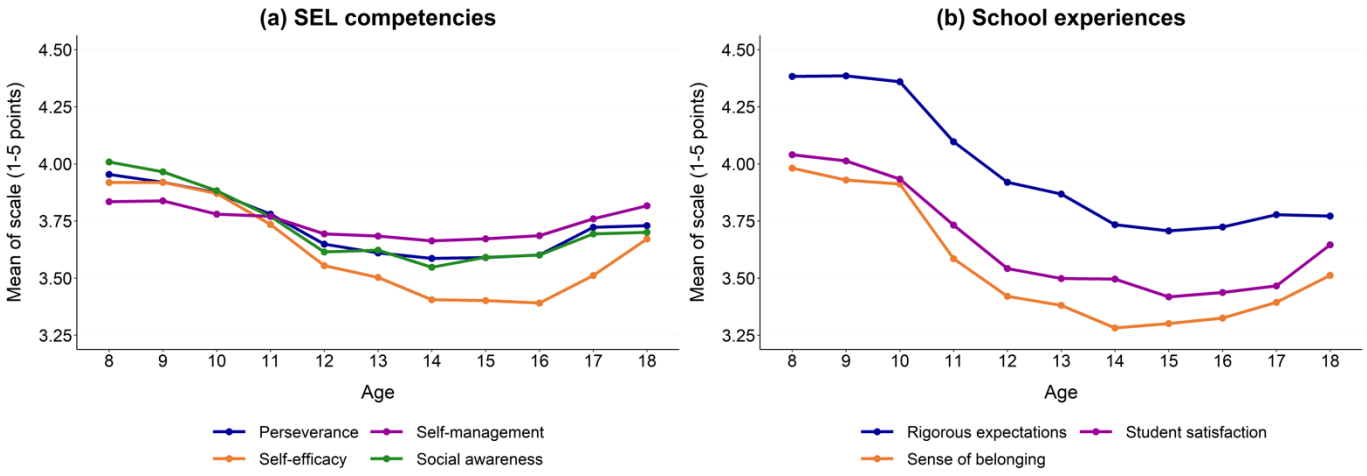
Appendix C. Supporting analysis

This appendix presents supporting analyses for each research question.

Research question 1. How do average social and emotional learning (SEL) competencies and school experiences differ across grade levels and change for individual students between years? *Do student and teacher reports of SEL competencies and school experiences change in similar ways across grade levels? To what extent do the average differences in students' SEL competencies and school experiences across grades differ by the type of students (such as students classified by gender, race/ethnicity, and academic achievement)? How are individual students' reports of SEL competencies and school experiences associated between years, and how does that association compare to that of other variables (such as achievement test scores, absences, and suspensions)?*

Students' average SEL competencies and school experiences exhibited similar patterns across ages as they did across grades (figure C1). As in the analyses across grades, the averages across ages for both SEL competencies and school experiences exhibited a U-shaped pattern. Some measures varied more than others. Self-management was relatively stable across age, whereas sense of belonging varied the most.

Figure C1. Students self-reported SEL competencies and school experiences exhibited similar patterns by age as by grade



SEL is social and emotional learning.

Note: The figure shows for each student ages 8–18 the mean of each SEL competency and school climate scale (described in appendix B). The means were calculated using nonresponse weights, as described in appendix B.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Students' SEL competencies and school experiences also showed similar patterns across subgroups of students as they do in the full sample (tables C1–C7). Across subgroups, students tended to have the lowest SEL competencies and school experiences during middle school (grades 6–8) and early high school (grades 9 and 10). However, as discussed in the main text, some subgroups tended to have higher levels of SEL competencies and school experiences for all grades.

Table C1. Average level of perseverance by subgroup and grade

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
All	3.95	3.93	3.87	3.77	3.63	3.61	3.56	3.59	3.65	3.79	<0.001
Gender											
Male	3.89	3.86	3.83	3.76	3.63	3.61	3.52	3.59	3.65	3.77	<0.001
Female	4.02	3.99	3.92	3.77	3.64	3.60	3.60	3.58	3.65	3.80	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.17***	0.17***	0.12***	0.01	0.01	0.01	0.10	0.01	0.00	0.04	
Survey year											
2017/18	3.95	3.91	3.88	3.78	3.65	3.63	3.50	3.56	3.63	3.80	<0.001
2018/19	3.96	3.94	3.87	3.75	3.62	3.59	3.62	3.62	3.67	3.77	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.01	0.04	0.01	0.04	0.04	0.05	0.16**	0.08	0.05	0.04	
Race/ethnicity											
Black	4.02	3.99	3.92	3.83	3.69	3.68	3.58	3.63	3.70	3.83	<0.001
Hispanic	4.00	3.96	3.88	3.73	3.60	3.54	3.52	3.49	3.57	3.72	<0.001
White	3.70	3.66	3.64	3.54	3.41	3.38	3.44	3.52	3.48	3.50	<0.001
Other	3.79	3.82	3.74	3.67	3.61	3.50	3.49	3.57	3.64	3.73	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.42***	0.43***	0.36***	0.38***	0.36***	0.39***	0.18**	0.18***	0.29***	0.43***	
At-risk status											
At risk	4.03	4.02	3.93	3.83	3.71	3.66	3.58	3.65	3.69	3.81	<0.001
Not at risk	3.89	3.85	3.82	3.71	3.58	3.58	3.53	3.52	3.62	3.77	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.18***	0.22***	0.14***	0.16***	0.17***	0.10**	0.07	0.17***	0.09	0.05	
Ward											
1	4.00	3.86	3.90	3.70	3.64	3.58	3.44	3.56	3.56	3.78	<0.001
2	3.95	3.79	3.74	3.69	3.60	3.65	3.66	3.57	3.63	3.59	<0.001
3	3.68	3.69	3.64	3.58	3.41	3.39	3.44	3.48	3.51	3.59	<0.001
4	4.04	3.99	3.82	3.83	3.74	3.67	3.64	3.64	3.70	3.75	<0.001
5	4.09	4.03	3.92	3.87	3.75	3.73	3.63	3.62	3.70	3.86	<0.001
6	3.89	3.89	3.91	3.84	3.70	3.65	3.59	3.51	3.61	3.78	<0.001
7	4.07	4.10	3.95	3.83	3.77	3.75	3.66	3.61	3.73	3.87	<0.001
8	4.02	3.99	4.01	3.83	3.67	3.69	3.54	3.71	3.71	3.75	<0.001

Table C1. Average level of perseverance by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Absolute value of the standardized difference in maximum and minimum value	0.53***	0.53***	0.48***	0.38***	0.47***	0.47***	0.29***	0.30**	0.29***	0.36***	
Special education status											
Special education	3.85	3.84	3.79	3.76	3.61	3.57	3.48	3.59	3.49	3.84	<0.001
Not special education	3.97	3.94	3.89	3.77	3.64	3.62	3.58	3.58	3.68	3.77	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.16***	0.13***	0.13***	0.01	0.04	0.07	0.13	0.01	0.25**	0.09	
English learner status											
English learner	4.00	3.99	3.85	3.74	3.73	3.60	3.56	3.53	3.57	3.74	<0.001
Not English learner	3.94	3.92	3.87	3.77	3.62	3.61	3.56	3.59	3.66	3.79	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.08*	0.09**	0.03	0.04	0.14*	0.01	0.00	0.08	0.12	0.07	
Proficient/college ready in math and ELA											
Proficient/college ready	3.88	3.83	3.83	3.66	3.56	3.56	3.63	3.68	3.61	3.67	<0.001
Not proficient/ college ready	3.99	3.97	3.89	3.79	3.66	3.64	3.59	3.58	3.66	3.80	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.14***	0.18***	0.08**	0.17***	0.13***	0.10**	0.05	0.13	0.07	0.17***	
Chronic absentee status											
Chronically absent	3.98	3.98	3.88	3.74	3.61	3.56	3.49	3.56	3.62	3.79	<0.001
Not chronically absent	3.95	3.92	3.87	3.77	3.64	3.63	3.65	3.63	3.71	3.77	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.04	0.08*	0.01	0.04	0.04	0.09*	0.21***	0.09*	0.12**	0.03	

Table C1. Average level of perseverance by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Suspension status											
Ever suspended	3.85	3.86	3.75	3.71	3.60	3.54	3.42	3.53	3.58	3.82	<0.001
Never suspended	3.96	3.93	3.88	3.78	3.64	3.63	3.59	3.60	3.66	3.78	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.14	0.09	0.17**	0.09	0.05	0.12**	0.22*	0.09	0.10	0.05	

Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts.

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Note: The table shows the average level of perseverance (described in appendix B) for each grade level, over both the full sample and for ten sets of subgroups. The means were calculated using nonresponse weights, as described in appendix B.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C2. Average level of rigorous expectations by subgroup and grade

Sample	Grade										*p*-value from test of equality across grades
	3	4	5	6	7	8	9	10	11	12	
All	4.37	4.39	4.37	4.06	3.90	3.86	3.72	3.69	3.76	3.84	<0.001
Gender											
Male	4.31	4.34	4.30	4.03	3.89	3.84	3.67	3.66	3.74	3.83	<0.001
Female	4.44	4.44	4.44	4.09	3.90	3.88	3.78	3.72	3.79	3.86	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.17***	0.13***	0.18***	0.08**	0.01	0.05	0.14**	0.08	0.06	0.04	
Survey year											
2017/18	4.37	4.36	4.34	4.04	3.88	3.82	3.65	3.63	3.72	3.81	<0.001
2018/19	4.38	4.42	4.39	4.07	3.91	3.90	3.79	3.74	3.80	3.88	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.01	0.08***	0.06**	0.04	0.04	0.10**	0.18**	0.14*	0.10*	0.09	
Race/ethnicity											
Black	4.41	4.41	4.39	4.07	3.88	3.87	3.71	3.68	3.77	3.87	<0.001
Hispanic	4.36	4.40	4.36	4.03	3.92	3.86	3.73	3.70	3.74	3.84	<0.001
White	4.29	4.32	4.27	4.01	3.92	3.82	3.74	3.68	3.72	3.61	<0.001
Other	4.24	4.28	4.30	4.11	4.00	3.89	3.77	3.75	3.82	3.79	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.22***	0.17***	0.15***	0.13	0.15*	0.09	0.08	0.09	0.13	0.33***	
At-risk status											
At risk	4.42	4.44	4.40	4.08	3.92	3.89	3.70	3.70	3.77	3.85	<0.001
Not at risk	4.34	4.35	4.34	4.04	3.88	3.85	3.74	3.67	3.75	3.84	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.10***	0.12***	0.08***	0.05*	0.05	0.05	0.05	0.04	0.03	0.01	
Ward											
1	4.36	4.30	4.34	3.99	3.98	3.91	3.69	3.76	3.76	3.88	<0.001
2	4.28	4.22	4.23	4.00	3.84	3.98	3.94	3.88	3.78	3.79	<0.001
3	4.25	4.26	4.27	4.04	3.90	3.84	3.66	3.52	3.59	3.55	<0.001
4	4.41	4.49	4.33	4.04	3.95	3.87	3.81	3.84	3.83	3.92	<0.001
5	4.41	4.46	4.37	4.04	3.92	3.85	3.69	3.53	3.63	3.85	<0.001
6	4.34	4.36	4.45	4.08	3.86	3.81	3.78	3.53	3.80	3.95	<0.001
7	4.49	4.50	4.39	4.13	3.88	3.90	3.64	3.67	3.77	3.72	<0.001
8	4.41	4.40	4.44	4.10	3.87	3.84	3.64	3.66	3.77	3.78	<0.001

Table C2. Average level of rigorous expectations by subgroup and grade (continued)

Sample	Grade										*p*-value from test of equality across grades
	3	4	5	6	7	8	9	10	11	12	
Absolute value of the standardized difference in maximum and minimum value	0.31***	0.36***	0.28***	0.18*	0.18*	0.22***	0.38***	0.46***	0.31**	0.51***	
Special education status											
Special education	4.26	4.28	4.27	3.96	3.85	3.75	3.62	3.64	3.63	3.86	<0.001
Not special education	4.39	4.41	4.39	4.08	3.91	3.88	3.75	3.70	3.79	3.84	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.17***	0.17***	0.15***	0.15**	0.08	0.17**	0.17	0.08	0.20*	0.03	
English learner status											
English learner	4.34	4.36	4.27	3.86	3.89	3.83	3.71	3.75	3.76	3.93	<0.001
Not English learner	4.38	4.39	4.38	4.08	3.90	3.86	3.72	3.68	3.76	3.84	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.05	0.04	0.14***	0.28***	0.01	0.04	0.01	0.09	0.00	0.12	
Proficient/college ready in math and ELA											
Proficient/college ready	4.40	4.42	4.41	4.13	4.00	3.94	3.92	3.92	3.82	3.85	<0.001
Not proficient/ college ready	4.37	4.38	4.36	4.04	3.86	3.85	3.71	3.68	3.74	3.83	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.04*	0.05*	0.06***	0.12***	0.18***	0.12**	0.27***	0.31***	0.10*	0.03	
Chronic absentee status											
Chronically absent	4.38	4.41	4.30	3.98	3.79	3.75	3.63	3.63	3.72	3.83	<0.001
Not chronically absent	4.37	4.39	4.38	4.08	3.94	3.91	3.85	3.79	3.86	3.88	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.01	0.03	0.10**	0.13**	0.19***	0.20***	0.28***	0.20***	0.18***	0.06	

Table C2. Average level of rigorous expectations by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Suspension status											
Ever suspended	4.36	4.30	4.31	3.93	3.74	3.71	3.48	3.62	3.68	3.78	<0.001
Never suspended	4.37	4.39	4.37	4.08	3.93	3.90	3.78	3.70	3.77	3.85	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.01	0.12*	0.08	0.19***	0.24***	0.24***	0.38***	0.10	0.12	0.09	

Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts.
* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Note: The table shows the average level of rigorous expectations (described in appendix B) for each grade level, over both the full sample and for ten sets of subgroups. The means were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C3. Average level of self-efficacy by subgroup and grade

Sample	Grade										*p*-value from test of equality across grades
	3	4	5	6	7	8	9	10	11	12	
All	3.92	3.92	3.87	3.72	3.53	3.50	3.40	3.40	3.45	3.61	<0.001
Gender											
Male	3.89	3.91	3.86	3.76	3.59	3.55	3.42	3.46	3.52	3.64	<0.001
Female	3.95	3.93	3.89	3.67	3.47	3.44	3.38	3.34	3.37	3.58	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.07**	0.02	0.04	0.11***	0.14***	0.13***	0.05	0.14**	0.18***	0.07	
Survey year											
2017/18	3.91	3.92	3.88	3.72	3.55	3.51	3.37	3.34	3.42	3.60	<0.001
2018/19	3.93	3.93	3.86	3.72	3.51	3.49	3.43	3.47	3.48	3.61	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.02	0.01	0.02	0.00	0.05	0.02	0.07	0.16**	0.07	0.01	
Race/ethnicity											
Black	3.95	3.93	3.88	3.73	3.53	3.50	3.39	3.43	3.48	3.66	<0.001
Hispanic	3.92	3.90	3.80	3.62	3.46	3.39	3.38	3.31	3.36	3.47	<0.001
White	3.83	3.96	3.97	3.81	3.61	3.64	3.52	3.46	3.46	3.48	<0.001
Other	3.80	3.87	3.79	3.74	3.55	3.53	3.41	3.37	3.34	3.50	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.18***	0.11	0.22***	0.23***	0.18***	0.30***	0.17**	0.18**	0.17*	0.23***	
At-risk status											
At risk	3.97	3.96	3.86	3.72	3.51	3.47	3.39	3.44	3.50	3.61	<0.001
Not at risk	3.88	3.90	3.88	3.72	3.54	3.52	3.41	3.36	3.40	3.61	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.11***	0.07**	0.02	0.00	0.04	0.06	0.02	0.10*	0.12*	0.00	
Ward											
1	3.91	3.77	3.86	3.58	3.55	3.50	3.27	3.37	3.38	3.56	<0.001
2	3.85	3.83	3.78	3.70	3.52	3.55	3.49	3.40	3.37	3.45	<0.001
3	3.75	3.89	3.82	3.69	3.49	3.47	3.42	3.29	3.28	3.36	<0.001
4	3.97	3.97	3.85	3.72	3.50	3.52	3.58	3.41	3.57	3.64	<0.001
5	3.94	3.92	3.79	3.77	3.59	3.55	3.37	3.38	3.50	3.66	<0.001
6	3.90	3.93	3.91	3.78	3.57	3.50	3.32	3.28	3.34	3.59	<0.001
7	4.09	4.02	3.92	3.72	3.55	3.55	3.43	3.47	3.51	3.57	<0.001
8	3.92	3.93	3.92	3.72	3.51	3.45	3.41	3.57	3.50	3.65	<0.001

Table C3. Average level of self-efficacy by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Absolute value of the standardized difference in maximum and minimum value	0.41***	0.30***	0.17**	0.24***	0.12	0.12	0.37***	0.35**	0.35***	0.36***	
Special education status											
Special education	3.76	3.77	3.71	3.67	3.43	3.43	3.32	3.36	3.32	3.63	<0.001
Not special education	3.95	3.95	3.91	3.73	3.55	3.51	3.42	3.41	3.48	3.60	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.23***	0.22***	0.24***	0.07	0.14**	0.10	0.12	0.06	0.19*	0.04	
English learner status											
English learner	3.91	3.88	3.72	3.55	3.53	3.41	3.39	3.42	3.42	3.54	<0.001
Not English learner	3.92	3.93	3.89	3.73	3.53	3.51	3.40	3.40	3.45	3.61	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.01	0.06	0.20***	0.22***	0.00	0.12*	0.01	0.02	0.04	0.08	
Proficient/college ready in math and ELA											
Proficient/college ready	3.98	4.06	4.07	3.86	3.72	3.74	3.60	3.42	3.50	3.55	<0.001
Not proficient/ college ready	3.89	3.86	3.81	3.67	3.46	3.44	3.37	3.35	3.41	3.58	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.11***	0.24***	0.31***	0.23***	0.31***	0.36***	0.27***	0.08	0.11*	0.04	
Chronic absentee status											
Chronically absent	3.91	3.89	3.77	3.65	3.42	3.38	3.33	3.39	3.43	3.62	<0.001
Not chronically absent	3.92	3.93	3.89	3.74	3.56	3.55	3.50	3.43	3.49	3.56	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.01	0.05	0.14***	0.11*	0.17***	0.20***	0.20***	0.05	0.07	0.07	

Table C3. Average level of self-efficacy by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Suspension status											
Ever suspended	3.83	3.75	3.75	3.65	3.49	3.45	3.30	3.45	3.41	3.68	<0.001
Never suspended	3.92	3.93	3.88	3.73	3.53	3.51	3.42	3.39	3.45	3.60	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.11	0.22**	0.16**	0.10	0.05	0.07	0.14	0.07	0.05	0.10	

Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts.
* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Note: The table shows the average level of self-efficacy scale (described in appendix B) for each grade level, over both the full sample and for ten sets of subgroups. The means were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C4. Average level of self-management by subgroup and grade

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
All	3.83	3.84	3.78	3.77	3.67	3.68	3.62	3.67	3.76	3.84	<0.001
Gender											
Male	3.73	3.74	3.69	3.73	3.65	3.65	3.55	3.62	3.70	3.80	<0.001
Female	3.93	3.95	3.87	3.81	3.70	3.71	3.69	3.72	3.82	3.88	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.27***	0.28***	0.24***	0.11**	0.07*	0.08*	0.19***	0.13**	0.16**	0.11*	
Survey year											
2017/18	3.80	3.81	3.76	3.75	3.71	3.66	3.58	3.63	3.74	3.86	<0.001
2018/19	3.86	3.86	3.79	3.78	3.64	3.69	3.65	3.71	3.78	3.82	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.08**	0.07**	0.04	0.04	0.09**	0.04	0.09	0.11*	0.05	0.05	
Race/ethnicity											
Black	3.76	3.75	3.70	3.70	3.60	3.62	3.57	3.65	3.74	3.84	<0.001
Hispanic	3.85	3.90	3.82	3.75	3.67	3.65	3.61	3.63	3.70	3.83	<0.001
White	4.02	4.09	4.05	4.05	3.96	3.94	3.89	3.88	3.96	3.88	<0.001
Other	3.91	3.97	3.97	4.00	3.90	3.85	3.91	3.81	3.90	3.79	0.084
Absolute value of the standardized difference in maximum and minimum value	0.35***	0.45***	0.47***	0.47***	0.48***	0.43***	0.45***	0.33***	0.35***	0.12	
At-risk status											
At risk	3.78	3.74	3.69	3.68	3.57	3.57	3.55	3.62	3.72	3.80	<0.001
Not at risk	3.86	3.91	3.86	3.84	3.74	3.75	3.68	3.72	3.79	3.87	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.11***	0.23***	0.23***	0.21***	0.23***	0.24***	0.17**	0.13*	0.09	0.09*	
Ward											
1	3.79	3.73	3.82	3.64	3.68	3.57	3.51	3.64	3.68	3.84	<0.001
2	3.79	3.84	3.68	3.86	3.73	3.80	3.93	3.93	3.87	3.84	<0.001
3	3.93	4.02	3.94	3.90	3.78	3.80	3.77	3.72	3.82	3.81	<0.001
4	3.89	3.92	3.82	3.76	3.72	3.68	3.57	3.62	3.68	3.81	<0.001
5	3.72	3.78	3.68	3.72	3.57	3.66	3.65	3.62	3.76	3.83	0.002
6	3.85	3.86	3.78	3.81	3.73	3.68	3.45	3.48	3.60	3.79	<0.001
7	3.82	3.78	3.69	3.72	3.60	3.60	3.51	3.64	3.80	3.89	<0.001
8	3.73	3.69	3.70	3.63	3.49	3.53	3.46	3.56	3.71	3.74	<0.001

Table C4. Average level of self-management by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Absolute value of the standardized difference in maximum and minimum value	0.28***	0.44***	0.35***	0.36***	0.39***	0.36***	0.64***	0.60***	0.36***	0.20	
Special education status											
Special education	3.63	3.61	3.60	3.61	3.51	3.47	3.41	3.47	3.50	3.79	<0.001
Not special education	3.86	3.88	3.82	3.80	3.71	3.72	3.67	3.72	3.81	3.85	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.31***	0.36***	0.29***	0.25***	0.27***	0.33***	0.35***	0.33***	0.41***	0.08	
English learner status											
English learner	3.81	3.82	3.70	3.65	3.66	3.61	3.52	3.63	3.63	3.81	<0.001
Not English learner	3.83	3.84	3.79	3.78	3.67	3.68	3.63	3.67	3.77	3.84	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.03	0.03	0.12**	0.17**	0.01	0.09	0.15	0.05	0.19**	0.04	
Proficient/college ready in math and ELA											
Proficient/ college ready	4.06	4.11	4.07	4.04	3.96	4.01	3.99	3.93	3.90	3.86	<0.001
Not proficient/ college ready	3.74	3.72	3.68	3.69	3.58	3.61	3.61	3.61	3.74	3.85	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.43***	0.52***	0.52***	0.47***	0.51***	0.53***	0.51***	0.43***	0.21***	0.01	
Chronic absentee status											
Chronically absent	3.73	3.71	3.62	3.59	3.45	3.44	3.45	3.57	3.69	3.81	<0.001
Not chronically absent	3.84	3.86	3.80	3.82	3.75	3.77	3.85	3.84	3.89	3.98	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.15***	0.20***	0.24***	0.31***	0.40***	0.44***	0.53***	0.36***	0.27***	0.23***	

Table C4. Average level of self-management by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Suspension status											
Ever suspended	3.38	3.36	3.36	3.44	3.36	3.38	3.33	3.43	3.44	3.67	0.001
Never suspended	3.84	3.86	3.81	3.82	3.74	3.75	3.69	3.72	3.81	3.86	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.61***	0.67***	0.60***	0.51***	0.51***	0.49***	0.48***	0.39***	0.49***	0.25**	

Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts.

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Note: The table shows the average level of self-management (described in appendix B) for each grade level, over both the full sample and for ten sets of subgroups. The means were calculated using nonresponse weights, as described in appendix B.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C5. Average level of sense of belonging by subgroup and grade

Sample	Grade										p-value from test of equality across grades
	3	4	5	6	7	8	9	10	11	12	
All	3.98	3.94	3.91	3.55	3.41	3.38	3.29	3.30	3.37	3.47	<0.001
Gender											
Male	3.93	3.92	3.91	3.57	3.48	3.45	3.32	3.37	3.43	3.53	<0.001
Female	4.02	3.96	3.92	3.52	3.33	3.30	3.26	3.22	3.30	3.40	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.11***	0.05*	0.01	0.06*	0.18***	0.18***	0.07	0.18***	0.15***	0.15**	
Survey year											
2017/18	3.95	3.92	3.90	3.55	3.42	3.37	3.29	3.29	3.38	3.47	<0.001
2018/19	4.00	3.95	3.92	3.55	3.39	3.39	3.29	3.32	3.36	3.46	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.06*	0.04	0.02	0.00	0.04	0.02	0.00	0.04	0.02	0.01	
Race/ethnicity											
Black	3.95	3.88	3.89	3.54	3.38	3.38	3.27	3.29	3.36	3.49	<0.001
Hispanic	4.07	4.07	3.99	3.60	3.50	3.40	3.32	3.32	3.38	3.39	<0.001
White	3.94	3.95	3.91	3.47	3.36	3.36	3.38	3.37	3.43	3.44	<0.001
Other	3.92	3.94	3.97	3.53	3.46	3.31	3.32	3.31	3.37	3.39	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.18***	0.22***	0.12***	0.15***	0.16***	0.11	0.13*	0.09	0.08	0.12*	
At-risk status											
At risk	3.99	3.92	3.90	3.56	3.41	3.40	3.30	3.31	3.39	3.45	<0.001
Not at risk	3.97	3.94	3.93	3.54	3.40	3.37	3.29	3.30	3.35	3.48	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.02	0.02	0.04	0.02	0.01	0.04	0.01	0.01	0.05	0.04	
Ward											
1	4.09	3.93	3.96	3.50	3.48	3.37	3.19	3.35	3.34	3.41	<0.001
2	3.99	3.79	3.82	3.44	3.38	3.35	3.52	3.47	3.40	3.37	<0.001
3	3.91	3.94	3.92	3.55	3.37	3.36	3.31	3.26	3.34	3.33	<0.001
4	4.04	4.06	3.95	3.66	3.51	3.44	3.35	3.27	3.25	3.45	<0.001
5	3.97	3.96	3.79	3.53	3.36	3.40	3.17	3.11	3.26	3.38	<0.001
6	3.94	3.90	3.95	3.54	3.41	3.35	3.23	3.21	3.26	3.37	<0.001
7	4.03	3.97	3.88	3.55	3.39	3.48	3.23	3.31	3.45	3.40	<0.001
8	3.91	3.82	3.91	3.50	3.35	3.33	3.35	3.19	3.41	3.47	<0.001

Table C5. Average level of sense of belonging by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	p-value from test of equality across grades
Absolute value of the standardized difference in maximum and minimum value	0.21***	0.32***	0.20**	0.26***	0.19**	0.18*	0.41***	0.42***	0.23*	0.16	
Special education status											
Special education	3.93	3.84	3.85	3.60	3.49	3.36	3.27	3.38	3.34	3.58	<0.001
Not special education	3.99	3.95	3.93	3.53	3.39	3.38	3.30	3.28	3.38	3.44	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.07	0.13***	0.09**	0.08	0.12*	0.02	0.04	0.12	0.05	0.16*	
English learner status											
English learner	4.08	4.06	3.93	3.62	3.55	3.41	3.29	3.32	3.44	3.46	<0.001
Not English learner	3.96	3.91	3.91	3.54	3.39	3.38	3.29	3.30	3.36	3.47	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.14***	0.18***	0.02	0.09	0.19***	0.04	0.00	0.02	0.09	0.01	
Proficient/college ready in math and ELA											
Proficient/college ready	4.03	4.00	3.97	3.53	3.41	3.42	3.43	3.38	3.41	3.42	<0.001
Not proficient/ college ready	3.96	3.91	3.90	3.55	3.41	3.37	3.27	3.27	3.33	3.43	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.08***	0.11***	0.08***	0.02	0.00	0.06	0.19***	0.13*	0.09*	0.01	
Chronic absentee status											
Chronically absent	3.93	3.88	3.81	3.49	3.30	3.27	3.24	3.28	3.35	3.47	<0.001
Not chronically absent	3.98	3.94	3.93	3.56	3.44	3.42	3.37	3.34	3.42	3.46	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.06	0.07*	0.14***	0.08*	0.16***	0.18***	0.15***	0.07	0.08*	0.01	

Table C5. Average level of sense of belonging by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	p-value from test of equality across grades
Suspension status											
Ever suspended	3.79	3.82	3.81	3.51	3.35	3.29	3.22	3.41	3.39	3.43	<0.001
Never suspended	3.98	3.94	3.92	3.55	3.42	3.40	3.31	3.28	3.37	3.47	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.22**	0.14*	0.13*	0.05	0.08	0.13**	0.11	0.15*	0.02	0.05	

Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts.
* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Note: The table shows the average level of sense of belonging (described in appendix B) for each grade level, over both the full sample and for ten sets of subgroups. The means were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C6. Average level of social awareness by subgroup and grade

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
All	4.01	3.96	3.88	3.75	3.60	3.61	3.50	3.58	3.67	3.76	<0.001
Gender											
Male	3.92	3.87	3.81	3.71	3.59	3.59	3.43	3.53	3.61	3.70	<0.001
Female	4.09	4.06	3.96	3.80	3.62	3.63	3.58	3.63	3.74	3.83	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.23***	0.26***	0.21***	0.12***	0.04	0.05	0.21***	0.14*	0.18***	0.18***	
Survey year											
2017/18	3.99	3.96	3.88	3.76	3.62	3.62	3.51	3.56	3.65	3.77	<0.001
2018/19	4.02	3.97	3.88	3.74	3.59	3.60	3.50	3.60	3.69	3.75	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.04*	0.01	0.00	0.03	0.04	0.03	0.01	0.05	0.05	0.03	
Race/ethnicity											
Black	3.99	3.93	3.84	3.71	3.54	3.55	3.44	3.54	3.64	3.74	<0.001
Hispanic	4.08	4.04	3.93	3.78	3.67	3.63	3.56	3.57	3.66	3.81	<0.001
White	3.99	4.01	3.99	3.92	3.79	3.84	3.82	3.85	3.88	3.88	<0.001
Other	3.97	3.95	3.94	3.78	3.70	3.67	3.68	3.75	3.82	3.81	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.15**	0.15***	0.21***	0.29***	0.34***	0.40***	0.52***	0.43***	0.33***	0.19***	
At-risk status											
At risk	4.01	3.95	3.83	3.70	3.54	3.54	3.42	3.53	3.66	3.72	<0.001
Not at risk	4.00	3.98	3.92	3.79	3.65	3.65	3.58	3.63	3.68	3.80	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.01	0.04	0.12***	0.12***	0.15***	0.15***	0.22***	0.14*	0.03	0.11*	
Ward											
1	4.06	3.89	3.94	3.56	3.58	3.50	3.43	3.56	3.58	3.81	<0.001
2	4.02	3.95	3.85	3.77	3.60	3.64	3.92	3.92	3.86	3.90	<0.001
3	3.96	3.99	3.96	3.85	3.69	3.74	3.69	3.71	3.75	3.76	<0.001
4	4.10	4.06	3.88	3.82	3.70	3.62	3.45	3.57	3.66	3.84	<0.001
5	4.03	4.00	3.82	3.77	3.59	3.61	3.51	3.45	3.62	3.68	<0.001
6	3.97	3.96	3.90	3.75	3.63	3.62	3.36	3.40	3.52	3.69	<0.001
7	4.02	4.01	3.79	3.67	3.53	3.51	3.25	3.44	3.60	3.71	<0.001
8	3.93	3.84	3.88	3.67	3.43	3.46	3.31	3.40	3.63	3.55	<0.001
Absolute value of the standardized difference in	0.23***	0.30***	0.23***	0.40***	0.37***	0.38***	0.92***	0.71***	0.47***	0.48***	

Table C6. Average level of social awareness by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
maximum and minimum value											
Special education status											
Special education	3.85	3.85	3.78	3.65	3.55	3.52	3.28	3.47	3.45	3.69	<0.001
Not special education	4.03	3.99	3.90	3.77	3.62	3.63	3.56	3.61	3.71	3.78	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.25***	0.19***	0.16***	0.16***	0.10	0.15**	0.38***	0.19*	0.36***	0.12	
English learner status											
English learner	4.06	4.01	3.86	3.71	3.69	3.56	3.48	3.58	3.62	3.83	<0.001
Not English learner	3.99	3.96	3.89	3.75	3.60	3.61	3.51	3.58	3.67	3.76	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.10**	0.07	0.04	0.05	0.12*	0.07	0.04	0.00	0.07	0.10	
Proficient/college ready in math and ELA											
Proficient/college ready	4.06	4.05	4.00	3.86	3.74	3.81	3.82	3.74	3.84	3.89	<0.001
Not proficient/ college ready	3.98	3.93	3.84	3.72	3.56	3.57	3.50	3.50	3.64	3.74	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.11***	0.16***	0.22***	0.19***	0.25***	0.33***	0.44***	0.33***	0.27***	0.21***	
Chronic absentee status											
Chronically absent	3.94	3.90	3.79	3.65	3.47	3.45	3.35	3.50	3.61	3.73	<0.001
Not chronically absent	4.02	3.98	3.90	3.78	3.65	3.67	3.72	3.71	3.79	3.91	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.11*	0.11**	0.15***	0.18***	0.25***	0.30***	0.51***	0.29***	0.25***	0.25***	

Table C6. Average level of social awareness by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Suspension status											
Ever suspended	3.67	3.65	3.56	3.57	3.39	3.39	3.20	3.43	3.49	3.63	<0.001
Never suspended	4.02	3.98	3.90	3.78	3.65	3.66	3.58	3.61	3.70	3.78	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.48***	0.45***	0.47***	0.29***	0.36***	0.37***	0.52***	0.25**	0.29**	0.21*	

Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts.
* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Note: The table shows the average level of social awareness (described in appendix B) for each grade level, over both the full sample and for 10 sets of subgroups. The means were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C7. Average level of student satisfaction by subgroup and grade

Sample	Grade										*p*-value from test of equality across grades
	3	4	5	6	7	8	9	10	11	12	
All	4.04	4.02	3.93	3.72	3.53	3.49	3.50	3.44	3.48	3.50	<0.001
Gender											
Male	3.98	3.98	3.89	3.72	3.57	3.52	3.53	3.49	3.54	3.56	<0.001
Female	4.10	4.05	3.97	3.71	3.49	3.44	3.48	3.39	3.41	3.45	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.16***	0.09***	0.11***	0.01	0.11***	0.11**	0.07	0.13**	0.17***	0.15**	
Survey year											
2017/18	4.03	3.99	3.93	3.73	3.57	3.47	3.52	3.42	3.49	3.50	<0.001
2018/19	4.04	4.04	3.93	3.71	3.49	3.50	3.49	3.47	3.47	3.50	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.01	0.07*	0.00	0.03	0.11***	0.04	0.04	0.07	0.03	0.00	
Race/ethnicity											
Black	3.95	3.92	3.85	3.64	3.44	3.39	3.43	3.37	3.43	3.47	<0.001
Hispanic	4.18	4.16	4.05	3.84	3.71	3.65	3.66	3.59	3.57	3.59	<0.001
White	4.17	4.17	4.12	3.88	3.64	3.68	3.74	3.64	3.63	3.58	<0.001
Other	4.06	4.09	4.04	3.79	3.58	3.60	3.62	3.54	3.55	3.52	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.30***	0.33***	0.36***	0.32***	0.36***	0.38***	0.41***	0.36***	0.26***	0.16**	
At-risk status											
At risk	3.98	3.96	3.86	3.67	3.48	3.40	3.45	3.44	3.48	3.51	<0.001
Not at risk	4.08	4.06	3.99	3.76	3.57	3.54	3.55	3.45	3.48	3.50	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.13***	0.13***	0.17***	0.12***	0.12***	0.18***	0.13**	0.01	0.00	0.01	
Ward											
1	4.10	4.02	4.03	3.61	3.62	3.50	3.46	3.44	3.43	3.46	<0.001
2	4.11	3.94	3.90	3.77	3.55	3.54	4.01	3.86	3.79	3.65	<0.001
3	4.16	4.15	4.09	3.90	3.63	3.67	3.55	3.44	3.48	3.45	<0.001
4	4.11	4.15	3.94	3.82	3.65	3.58	3.62	3.58	3.46	3.60	<0.001
5	4.05	3.99	3.85	3.64	3.47	3.40	3.38	3.24	3.36	3.38	<0.001
6	3.99	3.97	3.95	3.60	3.38	3.26	3.27	3.22	3.26	3.34	<0.001
7	3.96	4.00	3.77	3.54	3.32	3.52	3.41	3.39	3.36	3.25	<0.001
8	3.90	3.82	3.87	3.64	3.46	3.32	3.28	3.16	3.43	3.31	<0.001

Table C7. Average level of student satisfaction by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Absolute value of the standardized difference in maximum and minimum value	0.34***	0.44***	0.42***	0.48***	0.44***	0.54***	0.98***	0.92***	0.70***	0.53***	
Special education status											
Special education	3.96	3.94	3.87	3.68	3.56	3.45	3.38	3.42	3.41	3.53	<0.001
Not special education	4.05	4.03	3.94	3.73	3.52	3.49	3.54	3.45	3.49	3.50	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.12***	0.12***	0.09**	0.07	0.05	0.05	0.21**	0.04	0.11	0.04	
English learner status											
English learner	4.17	4.14	4.02	3.85	3.84	3.77	3.73	3.71	3.71	3.76	<0.001
Not English learner	4.01	3.99	3.92	3.71	3.50	3.46	3.48	3.40	3.45	3.48	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.21***	0.20***	0.13***	0.18***	0.45***	0.41***	0.33***	0.41***	0.34***	0.37***	
Proficient/college ready in math and ELA											
Proficient/college ready	4.15	4.15	4.06	3.87	3.63	3.66	3.72	3.49	3.60	3.49	<0.001
Not proficient/ college ready	3.99	3.96	3.88	3.67	3.49	3.44	3.44	3.39	3.41	3.45	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.21***	0.25***	0.24***	0.26***	0.18***	0.29***	0.37***	0.13*	0.25***	0.05	
Chronic absentee status											
Chronically absent	3.92	3.91	3.78	3.58	3.38	3.34	3.40	3.41	3.46	3.50	<0.001
Not chronically absent	4.06	4.03	3.95	3.76	3.58	3.54	3.66	3.50	3.52	3.51	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.18***	0.16***	0.22***	0.24***	0.26***	0.26***	0.34***	0.12**	0.08	0.01	

Table C7. Average level of student satisfaction by subgroup and grade (continued)

Sample	Grade 3	4	5	6	7	8	9	10	11	12	*p*-value from test of equality across grades
Suspension status											
Ever suspended	3.78	3.80	3.75	3.54	3.35	3.27	3.32	3.35	3.41	3.44	<0.001
Never suspended	4.05	4.03	3.94	3.75	3.57	3.54	3.55	3.46	3.49	3.51	<0.001
Absolute value of the standardized difference in maximum and minimum value	0.36***	0.30***	0.25***	0.28***	0.29***	0.36***	0.30***	0.15*	0.11	0.09	

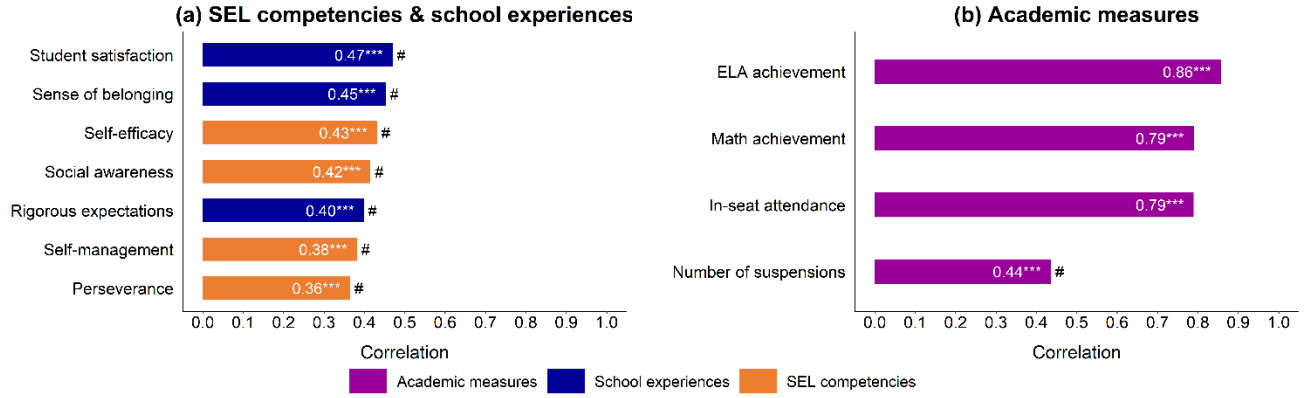
Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts.
* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Note: The table shows the average level of student satisfaction (described in appendix B) for each grade level, over both the full sample and for 10 sets of subgroups. The means were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

For individual students who took the surveys in consecutive years, the year-to-year correlations between SEL competencies and school experiences were high, ranging from 0.36 to 0.47 (figure C2). However, these estimates were much lower than the year-to-year correlations of all of the academic proficiency and behavior variables, except for the number of times students were suspended. For example, the year-to-year correlations in students' achievement test scores ranged from 0.79 to 0.86. Consistent with this finding, evidence has demonstrated that SEL competencies change more throughout the lifecycle than cognitive test scores, which stabilize earlier (Almlund et al., 2011). SEL competencies might be less stable because they reflect students' behavior, which could also change with students' circumstances (Heckman & Kautz, 2012).

One question is whether the lower year-to-year correlations occur because they are measured with more error. Although this study does not have the data to investigate this possibility, other evidence suggests that measurement error is unlikely to account for most of the differences in year-to-year correlations between academic measures and SEL competencies and school experiences. A study of the CORE districts found that higher levels of measurement error of SEL competencies accounted for only a small part of the difference in their year-to-year correlations relative to achievement test scores (West et al., 2018). In addition, other research has shown that younger students are more susceptible to acquiescence bias—the tendency to report positive values on surveys (Soto et al., 2008). Acquiescence bias might explain why elementary school students report higher values of SEL competencies and school experiences than middle school students. However, the teacher reports of the measures also decline across grades, suggesting that acquiescence bias does not likely account for all of the differences across grades. Reference bias could also arise if students rate themselves using different benchmarks (Duckworth & Yeager, 2015). Further study could explore these measurement issues.

Figure C2. Students' SEL competencies and school experiences were correlated between years but less so than most academic measures



ELA is English language arts. SEL is social and emotional learning.
*** Significant at $p < .001$.
# Correlation differs from ELA achievement correlation by 0.10 or more—the cutoff for a moderate difference between correlations.
Note: The figure shows the correlation between the 2017/18 and 2018/19 student reports of each SEL competency and school climate scale (described in appendix B), as well as between 2017/18 and 2018/19 data on four student academic measures. The correlations were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

The year-to-year correlations of students' SEL competencies and school experiences vary across subgroups of students, with differences ranging from low (0.00–0.09) to substantive (0.20–0.29) (table C8). For example, for each measure, the year-to-year correlations are lower by at least 0.02 for male students compared to female students. Similarly, the correlations are lower for students who receive special education services, are English learner students, are Hispanic or Black, and are not proficient/college ready in math and English language arts (ELA) than their subgroup counterparts. These differences in year-to-year correlations may reflect real differences in how some groups change, but they may also reflect measurement error. For example, students in some groups might find it challenging to understand items in the student survey or might select answers with less care.

Table C8. Year-to-year correlations of SEL competencies and school experiences

Sample	Perseverance	Self-management	Self-efficacy	Social awareness	Rigorous expectations	Sense of belonging	Student satisfaction
All	0.36	0.38	0.43	0.42	0.40	0.45	0.47
Gender							
Male	0.31	0.34	0.39	0.37	0.36	0.44	0.42
Female	0.42	0.42	0.48	0.46	0.45	0.46	0.52
Absolute value of the difference in maximum and minimum value	0.11	0.08	0.09	0.09	0.09	0.02	0.10
Race/ethnicity							
Black	0.34	0.35	0.42	0.39	0.39	0.44	0.46
Hispanic	0.35	0.39	0.41	0.44	0.37	0.45	0.41
White	0.47	0.50	0.54	0.48	0.50	0.52	0.53
Other	0.46	0.42	0.48	0.55	0.49	0.51	0.51
Absolute value of the difference in maximum and minimum value	0.13	0.15	0.13	0.16	0.13	0.08	0.12
At-risk status							
At risk	0.35	0.33	0.40	0.38	0.38	0.42	0.42
Not at risk	0.38	0.42	0.46	0.44	0.41	0.48	0.51
Absolute value of the difference in maximum and minimum value	0.03	0.09	0.06	0.06	0.03	0.06	0.09
Ward							
1	0.22	0.37	0.39	0.40	0.31	0.43	0.44
2	0.48	0.47	0.52	0.50	0.47	0.54	0.49
3	0.42	0.50	0.52	0.51	0.48	0.53	0.53
4	0.38	0.41	0.43	0.44	0.39	0.44	0.47
5	0.32	0.26	0.26	0.39	0.32	0.39	0.36
6	0.40	0.39	0.46	0.43	0.45	0.48	0.51
7	0.38	0.29	0.39	0.35	0.42	0.42	0.43
8	0.30	0.34	0.42	0.36	0.38	0.41	0.41
Absolute value of the difference in maximum and minimum value	0.26	0.24	0.26	0.16	0.17	0.15	0.17
Special education status							
Special education	0.26	0.24	0.33	0.30	0.24	0.35	0.37
Not special education	0.40	0.42	0.46	0.45	0.45	0.48	0.50
Absolute value of the difference in maximum and minimum value	0.14	0.18	0.13	0.15	0.21	0.13	0.13
English learner status							
English learner	0.29	0.34	0.33	0.37	0.32	0.43	0.35
Not English learner	0.38	0.39	0.44	0.42	0.41	0.45	0.48
Absolute value of the difference in maximum and minimum value	0.09	0.05	0.11	0.05	0.09	0.02	0.13
Proficient/college ready in math and ELA							
Proficient/college ready	0.47	0.51	0.52	0.49	0.47	0.51	0.52
Not proficient/college ready	0.37	0.37	0.43	0.40	0.41	0.45	0.47

Table C8. Year-to-year correlations of SEL competencies and school experiences (continued)

Sample	Perseverance	Self-management	Self-efficacy	Social awareness	Rigorous expectations	Sense of belonging	Student satisfaction
Absolute value of the difference in maximum and minimum value	0.10	0.14	0.09	0.09	0.06	0.06	0.05
Chronic absentee status							
Chronically absent	0.28	0.30	0.38	0.38	0.34	0.45	0.43
Not chronically absent	0.41	0.41	0.46	0.42	0.42	0.45	0.48
Absolute value of the difference in maximum and minimum value	0.13	0.11	0.08	0.04	0.08	0.00	0.05
Suspension status							
Ever suspended	0.18	0.15	0.28	0.30	0.23	0.39	0.41
Never suspended	0.40	0.41	0.46	0.42	0.42	0.46	0.47
Absolute value of the difference in maximum and minimum value	0.22	0.26	0.18	0.12	0.19	0.07	0.06

Absolute value of the difference between the maximum and minimum correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts. SEL is social and emotional learning.
Note: The table shows the pairwise correlation between the 2017/18 and 2018/19 versions of each SEL competency and school climate scale (described in appendix B) for the full sample and for nine sets of subgroups. The correlations were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Research question 2. To what extent do year-to-year changes in individual students' SEL competencies and school experiences differ across schools?

To help benchmark the difference in changes across schools, the study team calculated the average year-to-year changes in measures for average and high positive-change schools (those that are one standard deviation above average).[25] The difference in changes between schools was comparable to or greater than the changes within either type of school. Consistent with the analyses showing that SEL competencies and school experiences decline across some grades, the year-to-year changes were slightly negative for average change schools in some cases (table C9). In contrast, for high positive-change schools, the changes were always positive.

[25] Approximately 16 percent of schools would be at least one standard deviation above average.

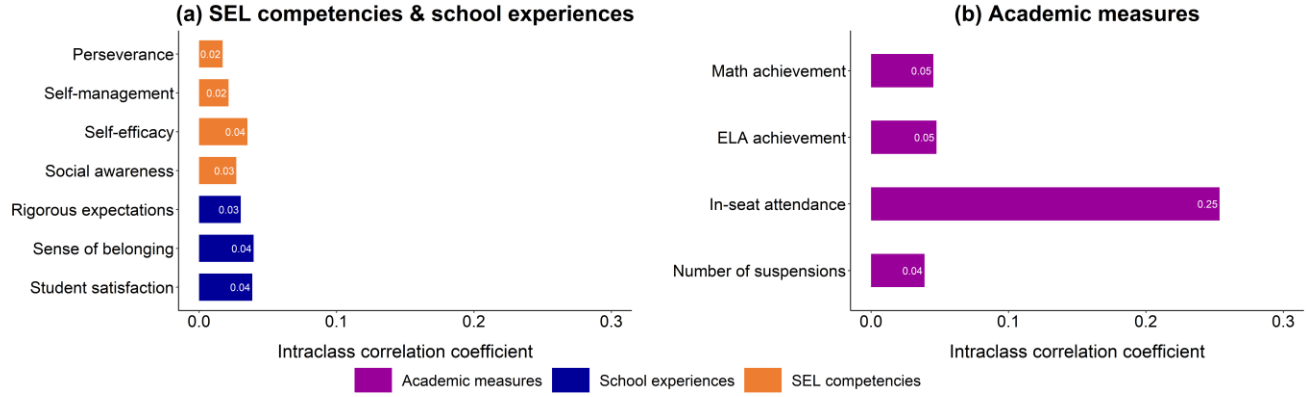
Table C9. Average year-to-year changes in measures for average and high positive-change schools

Measure	Outcome units	Outcome units			Percentile units		
		Average change school	High positive-change school	Difference	Average change school	High positive-change school	Difference
Perseverance	1- to 5-point scale	-0.03	0.08	0.11	-1.41	4.45	5.86
Self-management	1- to 5-point scale	0.02	0.13	0.12	0.83	7.31	6.47
Self-efficacy	1- to 5-point scale	-0.02	0.15	0.17	-0.99	6.94	7.93
Social awareness	1- to 5-point scale	-0.02	0.11	0.13	-0.92	6.15	7.07
Rigorous expectations	1- to 5-point scale	0.01	0.16	0.15	0.37	7.95	7.58
Sense of belonging	1- to 5-point scale	-0.03	0.15	0.18	-1.39	6.87	8.26
Student satisfaction	1- to 5-point scale	-0.07	0.08	0.15	-3.89	4.14	8.03
Math achievement	Standard deviations	0.00	0.13	0.13	0.00	5.50	5.49
ELA achievement	Standard deviations	-0.01	0.11	0.11	-0.29	4.36	4.65
In-seat attendance	Fraction	-0.03	0.08	0.11	-3.50	9.46	12.96
Number of suspensions	Number	-0.12	0.08	0.19	-4.84	3.11	7.94

ELA is English language arts.
Note: This table shows the average year-to-year changes in measures for average change schools and high change schools (those that are one standard deviation above average). The values were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

The intraclass correlation coefficients (ICCs) suggest similar conclusions about the extent to which year-to-year changes in the measures differ across schools as the metric presented in the main text (figure C3). Consistent with the findings presented in the main text, the ICC calculations suggest that changes in in-seat attendance vary the most across schools. However, changes in the SEL competencies and school experiences have similar but somewhat lower ICCs than changes in math and ELA achievement, which is the reverse of the percentile units presented in the main text. This difference arises because math and ELA achievement have higher year-to-year correlations, which reduces the differences in percentile units (see appendix B).

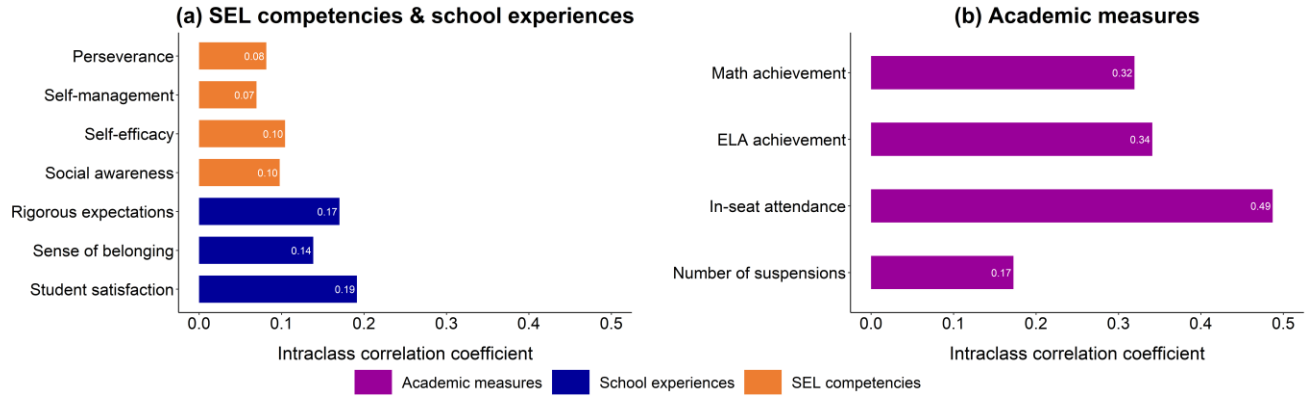
Figure C3. The ICCs for year-to-year changes in measures suggest similar conclusions as the percentile differences presented in the main text



ELA is English language arts. ICC is intraclass correlation coefficient. SEL is social and emotional learning.
Note: The figure shows the ICC of year-to-year changes in the measures across schools in 2017/18 (described in appendix B). The ICCs were calculated using nonresponse weights, as described in appendix B. The measures are defined in table B3.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18.

Compared to SEL competencies and school experiences, academic measures tend to vary more across schools, rather than within schools (figure C4). The ICCs for math achievement, ELA achievement, and in-seat attendance ranged from 0.32 to 0.49, whereas those for SEL competencies and school experiences ranged from 0.07 to 0.19. This finding is consistent with other research that has demonstrated that behavioral measures tend to have lower ICCs than achievement test scores (Schochet, 2008).

Figure C4. The ICCs for most academic measures exceeded those of SEL competencies and school experiences



ELA is English language arts. ICC is intraclass correlation coefficient. SEL is social and emotional learning.
Note: The figure shows the ICC of the measures across schools in 2017/18 (described in appendix B). The ICCs were calculated using nonresponse weights, as described in appendix B. The measures are defined in table B3.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18.

Research question 3. How do measures of SEL competencies and school experiences relate to future outcomes, and how do they complement other available data for predicting future outcomes? *To what extent do individual SEL competencies and school experiences relate to student outcomes measured one and two years later (such as achievement test scores, absences, suspensions, and whether a student feels loved, challenged, and prepared)? When other data are available—such as demographic information, achievement test scores, absences, and suspensions—to predict students' future outcomes, to what extent does adding measures of SEL competencies and school experiences improve the predictive power and accuracy of those predictions? Which types of data and statistical models could best help DCPS classify whether students are at risk of having negative outcomes?*

The relationship between grade-3 SEL competencies and school experiences and grade-3 achievement in ELA is similar to findings for the full sample (table C10). As with the full sample, self-management is the most related to achievement in ELA and perseverance is slightly negatively related to achievement in ELA.

Table C10. Pairwise correlations between grade-3 SEL competencies and school experiences and grade-3 ELA achievement

Achievement measure	Perseverance	Self-management	Self-efficacy	Social awareness	Rigorous Expectations	Sense of belonging	Student satisfaction
Grade-3 PARCC ELA Score (2017/18)	-0.01	0.22	0.09	0.13	0.10	0.08	0.17
Grade-3 PARCC ELA Score (2018/19)	-0.06	0.17	0.01	0.03	0.03	0.01	0.08

Absolute value of the correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts. PARCC is Partnership for Assessment of Readiness for College and Careers. SEL is social and emotional learning.
Note: This table shows the pairwise correlation between grade-3 PARCC ELA scores and individual SEL competencies and school experiences in 2017/18 and 2018/19 (described in appendix B). The correlations were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

SEL competencies and school experiences in grade 8 are related to measures of successful transitions to grade 9 (table C11). All SEL competencies and school experiences are negatively correlated with credits behind in grade 9 and positively correlated with successful grade progression. As in the full sample of students, self-management is the most related to successful outcomes.

Table C11. Pairwise correlations between grade-8 SEL competencies and school experiences in 2017/18 and grade-9 outcomes in 2018/19

Grade-9 outcome	Perseverance	Self-management	Self-efficacy	Social awareness	Rigorous expectations	Sense of belonging	Student satisfaction
Credits behind	-0.05	-0.14	-0.13	-0.11	-0.12	-0.03	-0.07
Progressed successfully	0.02	0.07	0.00	0.05	0.02	-0.03	0.01

Absolute value of the correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

SEL is social and emotional learning.
Note: This table shows the pairwise correlation between grade-9 outcomes in 2018/19 and individual SEL competencies and school experiences in 2017/18 (described in appendix B). The correlations were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

The study team estimated that credits behind in grade 9 accurately classifies eventual high school graduation 75 percent of the time (table C12). The regression coefficients between credits behind in successive years were positive and close to 1 (0.98 and 1.02), demonstrating that students who are more behind in one year are also likely to be more behind in the next year. Credits behind was negatively associated with graduating from high school, indicating that students who were behind were less likely to graduate.

Table C12. Results from analyses of the accuracy of classifying high school graduation on credits behind in grade 9

Description of estimate	Value of estimate
Regression of grade-10 credits behind on grade-9 credits behind	
Intercept	0.30
Coefficient	1.02
Regression of grade-11 credits behind on grade-10 credits behind	
Intercept	0.28
Coefficient	0.98
Probit model of high school graduation as a function of grade-11 credits behind	
Intercept	1.20
Coefficient	-0.32
Implied probit model of high school graduation as a function of grade-9 credits behind	
Intercept	0.97
Coefficient	-0.30
Estimated classification accuracy of grade-9 credits behind in predicting high school graduation based on simulation	0.75

Note: This table shows the output from a series of regression analyses designed to establish the relationship between credits behind in grade 9 and the probability of graduating high school (described in appendix B).
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

The correlations between predictors and outcomes one year out were similar to those between predictors and outcomes two years out (compare table C13 and table 1 in the main text). Compared to outcomes measured two years out, the correlations measured one year out are somewhat lower. However, the patterns of which variables are most predictive are similar, regardless of when the outcomes were measured.

Table C13. Correlations between predictors and students' outcomes measured one year out

Predictor(s) in 2017/18	Outcomes measured one year out in 2018/19: Progressed successfully (grades 3–12)	Number of suspensions (grades 3–12)	Math achievement (grades 3–12)	ELA achievement (grades 3–12)	In-seat attendance (grades 3–12)	Number of AP credits earned (grades 10–12)	Credits behind (grades9–12)	Graduated within 1 year (grades 12)	Loved, challenged, and prepared (grades 3-12)
(a) SEL competencies and climate perceptions									
Perseverance	0.05	0.04	-0.02	-0.03	0.09	0.00	0.08	0.04	0.28
Self-management	0.05	**0.13**	**0.24**	**0.25**	0.11	**0.17**	**0.10**	**0.13**	0.24
Self-efficacy	0.08	0.04	0.12	0.09	0.14	0.06	0.07	0.07	0.32
Social awareness	0.07	0.11	0.12	0.13	0.16	0.13	0.07	0.08	0.30
Rigorous expectations	**0.10**	0.06	0.07	0.07	**0.20**	0.03	0.05	0.04	0.30
Sense of belonging	0.08	0.04	0.05	0.04	0.16	0.01	0.03	0.05	**0.38**
Student satisfaction	0.07	0.09	0.13	0.12	0.17	0.05	0.03	-0.02	0.32
(b) Academic measures									
Math achievement	0.16	0.17	**0.83**	0.76	0.28	0.53	0.23	0.27	0.04
ELA achievement	0.17	0.17	0.75	**0.86**	0.29	**0.58**	0.27	0.30	0.03
In-seat attendance	**0.54**	0.13	0.28	0.28	**0.77**	0.33	**0.53**	**0.48**	**0.06**
Number of suspensions	-0.15	**-0.40**	-0.18	-0.20	-0.31	-0.17	-0.28	-0.21	-0.05
(c) Groups of variables									
SEL competencies and school climate perceptions	0.11	0.15	0.30	0.31	0.23	0.20	0.10	0.15	**0.42**
Demographics	0.29	0.24	0.59	0.65	0.56	0.54	0.27	0.29	0.25
Academic measures	**0.54**	**0.41**	**0.85**	**0.88**	**0.78**	**0.60**	**0.54**	**0.49**	0.07
(d) Combinations of groups of predictors									
Demographics and academic measures	0.56	0.43	0.86	0.89	0.80	0.67	0.61	0.51	0.26
All predictors	**0.56**	**0.44**	**0.86**	**0.89**	**0.80**	**0.67**	**0.62**	**0.52**	**0.44**

Absolute value of the correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

AP is Advanced Placement. ELA is English language arts. SEL is social and emotional learning.

Note: The table shows pairwise correlations and multivariate correlations between predictors in the left column and the outcomes in the top row. The outcomes were all recoded so that a higher value of the outcome is beneficial. The bold font indicates the correlation with the highest absolute value within the column and panel. The correlations were calculated using nonresponse weights, as described in appendix B. The sample included students who completed the Panorama survey in 2017/18.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

The study team examined the correlations between school-level teacher and parent reports of perseverance and rigorous expectations and school-level outcomes one year later (table C14). Both perseverance and rigorous expectations were correlated with outcomes, but the relative strength of the correlations differed by respondent type. The teacher reports of perseverance were more positively correlated with outcomes than were the reports of rigorous expectations. However, the reverse was true of the parent reports for academic behaviors (suspensions and attendance) and whether students felt loved, challenged, and prepared. Overall, the teacher reports tended to be the most predictive.

Table C14. Correlations between school-level parent and teacher reports of perseverance and rigorous expectations and school-level outcomes measured one year out

School-level predictor(s) in 2017/18	School-level outcomes measured one year out in 2018/19					
	Progressed successfully	Number of suspensions	Math achievement	ELA achievement	In-seat attendance	Loved, challenged, and prepared
Teacher reports						
Perseverance	0.37	0.56	0.52	0.54	0.54	0.56
Rigorous expectations	0.31	0.29	0.09	0.11	0.35	0.47
Parent reports						
Perseverance	-0.09	0.04	-0.11	-0.11	0.07	0.01
Rigorous expectations	0.09	0.38	0.01	0.01	0.36	0.44

Absolute value of the correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts.
Note: The table shows pairwise correlations between school-level predictors in the left column and school-level outcomes in the top row. The bold font indicates the correlation with the highest absolute value within the column. The school-level averages for teacher reports (but not parent reports) were calculated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

The classification accuracy for individual predictors and groups of predictors were consistent with the pairwise correlations and multivariate correlations presented in the main text (table C15). For all academic outcomes, individual academic measures accurately classified outcomes a higher percent of the time compared to individual SEL competencies or school experiences (compare panels a and b of table C15). Although the group of SEL competencies and school experiences improved the classification of academic outcomes relative to chance, they did so less than the group of demographic variables and academic measures (see panel c of table C15). Complex machine learning algorithms did not systematically perform better compared to probit models (compare panels c and d of table C15). Consistent with the multivariate correlations, predictive models that additionally included students' SEL competencies and school experiences improved the overall accuracy little beyond the models with students' demographic characteristics and academic measures (compare the last two rows in panels c and d of table C15). However, adding the SEL competencies and school experiences increased the accuracy by up to 7.8 percentage points when classifying whether students felt loved, challenged, and prepared in the future.

Table C15. Classification accuracy of predictors

Predictor(s)	Outcomes measured one year out in 2018/19: Did not progress successfully	Suspended	Not proficient/college ready in math	Not proficient/college ready in ELA	Chronically absent	Did not earn AP credits	Red grade progression status	Did not graduate	Not loved, challenged, and prepared	Outcomes measured two years out in 2019/20: Did not progress successfully	Suspended	Not college ready in math	Not college ready in ELA	Chronically absent	Did not earn AP credits	Red grade progression status	Did not graduate	Not loved, challenged, and prepared
Accuracy from assigning students to the most prevalent category	95.7	89.7	71.1	58.5	63.4	70.1	93.6	85.9	54.5	93.3	92.2	72.9	60.0	65.1	70.3	94.1	82.9	57.2
a. Improvements in accuracy using SEL competencies and school experiences (probit)																		
Perseverance	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	3.5
Self-management	0.0	0.0	0.0	1.7	0.1	0.0	0.0	0.0	7.1	0.0	0.0	0.0	1.2	0.1	0.0	0.0	0.0	3.7
Self-efficacy	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	9.9	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	5.1
Social awareness	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	9.7	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.4	5.3
Rigorous expectations	0.0	0.0	0.0	0.0	1.6	0.0	0.0	0.0	9.5	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.2	3.1
Sense of belonging	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	12.7	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	6.7
Student satisfaction	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	10.2	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	5.0
b. Improvements in accuracy using academic measures (probit)																		
Math achievement	0.0	0.0	15.5	22.9	2.0	4.7	0.0	0.4	0.1	0.0	0.0	14.9	22.3	1.1	3.4	0.0	0.0	0.1
ELA achievement	0.0	0.0	12.1	27.9	2.6	5.9	0.0	0.8	0.4	0.0	0.0	13.4	25.6	1.2	4.5	0.0	0.0	0.2
In-seat attendance	0.8	0.0	0.0	4.1	19.3	0.0	0.3	2.1	0.0	1.3	0.1	0.7	4.3	13.3	0.0	0.7	4.1	0.3
Number of suspensions	0.0	0.3	0.0	0.0	4.5	0.0	0.1	0.3	0.0	0.0	0.1	0.0	0.0	4.1	0.0	0.0	0.7	0.0
c. Improvements in accuracy using groups of variables (probit)																		
SEL competencies and school experiences	0.0	0.0	1.2	4.8	2.3	0.0	0.0	0.0	15.0	0.0	0.0	0.8	4.9	1.3	0.0	0.0	0.0	8.9
Basic demographics	0.0	0.0	9.7	16.9	14.0	8.2	0.0	0.1	8.0	0.9	0.0	8.6	16.4	9.0	9.4	0.0	0.2	5.7
Academic measures	0.9	0.3	16.4	28.8	19.2	5.9	0.5	2.4	0.1	1.5	0.2	16.3	26.3	13.2	4.3	0.1	4.6	0.0
Basic demographics and academic measures	1.2	0.4	18.5	29.7	21.6	13.9	1.8	3.5	8.0	2.0	0.2	17.6	27.7	15.7	14.8	0.3	5.2	5.9
All	1.3	0.4	18.6	29.7	21.7	14.0	2.0	3.7	15.8	2.0	0.1	17.6	27.8	15.7	14.8	0.6	5.8	11.1
d. Improvements in accuracy using groups of variables (random forests)																		
SEL competencies and school experiences	0.0	-0.1	1.6	5.7	1.9	-2.0	0.0	-0.3	12.9	-0.1	0.0	0.8	6.0	0.6	-1.3	0.0	-1.8	5.7
Basic demographics	0.0	0.0	9.8	16.5	12.5	7.8	0.0	0.0	8.0	0.7	0.0	8.5	15.5	8.7	9.7	0.0	0.0	5.4
Academic measures	0.1	-0.4	15.7	27.7	16.7	5.6	-0.3	-0.6	-0.6	0.3	-0.3	14.6	25.5	11.6	5.8	-0.2	0.4	-2.0
Basic demographics and academic measures	0.5	0.3	18.1	29.4	20.4	14.0	-0.1	0.1	7.6	1.4	0.2	17.0	27.1	15.0	14.8	-0.1	1.9	6.1
All	0.2	0.3	18.2	29.4	20.4	14.0	-0.2	1.6	14.2	1.2	0.1	16.6	27.3	14.8	14.7	-0.1	1.3	9.0

AP is Advanced Placement. ELA is English language arts. SEL is social and emotional learning.

Note: Baseline accuracy is the accuracy achieved by predicting the most frequently occurring outcome for each student. The subsequent rows show the improvements in accuracy by using each predictor or group of predictor variables (see appendix B).

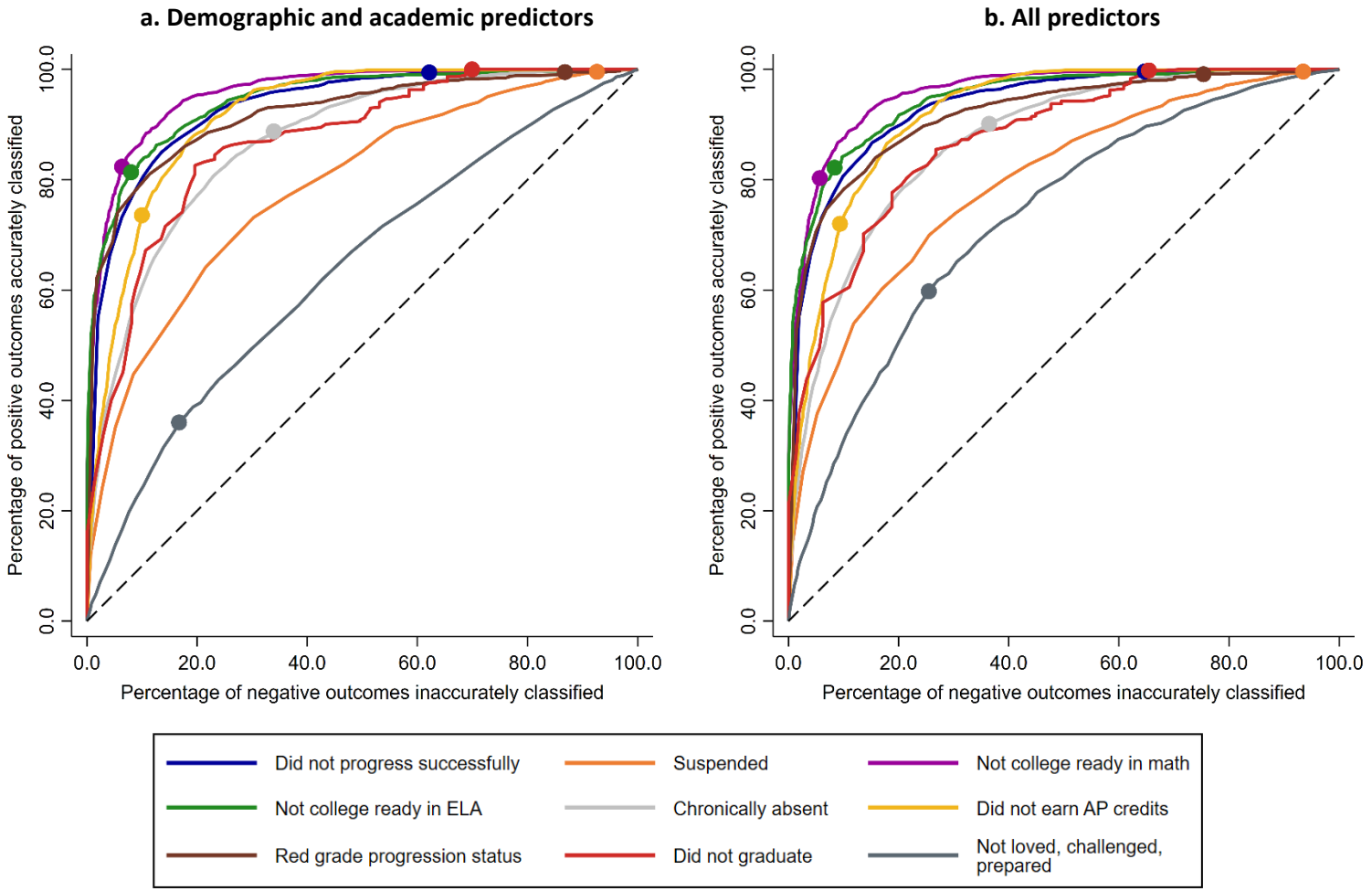
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20

When selecting a classification threshold, a receiver operating characteristic (ROC) curve can illustrate the tradeoff between accurately classifying students who eventually have positive outcomes (for example, graduating from high school) and inaccurately classifying those who eventually have negative outcomes (figure C5). For each model, each curve represents how the accuracy for these two groups of students changes as the threshold moves from 0 to 1. The dots on each curve represent the threshold that maximizes overall classification accuracy, which corresponds to the models used in this study. However, there are many potential decision rules to determine an optimal threshold for classifying observations.

Any potential decision rule involves tradeoffs that should be evaluated in the context of the model's intended usage. Reducing the threshold means that students who have a positive outcome are accurately classified a higher percentage of the time, but students with a negative outcome are accurately classified a lower percentage of the time. Depending on the relative importance of correctly or incorrectly classifying students with positive and negative outcomes, a different threshold may be appropriate. For example, for a model using demographic and academic predictors and a threshold that maximizes *total* accuracy (0.55), approximately 88.8 percent of students who are not chronically absent are classified accurately, and 66.1 percent of students who are chronically absent are classified accurately (see figure C5). However, if DCPS placed a higher weight on accurately classifying students who are eventually chronically absent compared to those who are not (rather than total accuracy), then a different threshold might be more appropriate. For example, a threshold of 0.90 would lead to a classification accuracy of 97.2 for those who are chronically absent but a classification accuracy of 34.2 for those who are not chronically absent.

Consistent with figure 4, these findings suggest that the SEL competencies and school experience variables add little to the ability to accurately classify students who eventually have positive or negative outcomes across all possible thresholds (compare panels a and b of figure C5). The one exception is classifying whether students felt loved, challenged, and prepared in the future, a measure based on future SEL competencies and school experiences. For this outcome, the ability to accurately classify students with both negative and positive outcomes generally improves with the inclusion of these additional predictors.

Figure C5. A ROC curve illustrates how a model's classification threshold governs the tradeoff between accurately classifying students with positive outcomes and misclassifying those with negative outcomes



AP is Advanced Placement. ELA is English language arts.

Note: The figure shows the ROC curves for outcomes measured two years out using the probit model with the sets of predictors listed above the figures. The dots represent the threshold that maximizes overall classification accuracy. The diagonal, dashed line represents a random classifier. A point in the upper left-hand corner represents perfect prediction.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2019/20.

Research question 4. How do measures of perseverance and rigorous expectations align across students, teachers, and parents? *Across schools, to what extent do survey reports on these measures from students, parents, and teachers align? Is alignment associated with characteristics of schools (such as the demographic characteristics of their student population and the schools' accountability ratings) and response rates on the survey?*

The alignment between student, parent, and teacher reports of perseverance and rigorous expectations differed across types of schools (tables C16–C19). In general, alignment was the lowest in schools with low School Transparency and Reporting Framework (STAR) ratings, a high percentage of Black students, and low student survey completion rates. The alignment also differed by geographic ward.

Table C16. Difference between respondent types overall and by school characteristics for perseverance

	Teacher-student		Teacher-parent		Student-parent	
Sample	Difference in scale	Standardized difference	Difference in scale	Standardized difference	Difference in scale	Standardized difference
All	-0.67***	-0.87	-0.54***	-0.70	0.12***	0.16
Ward						
1	-0.60***	-0.78	-0.56***	-0.73	0.04	0.05
2	-0.43***	-0.56	-0.40***	-0.52	0.03	0.04
3	-0.19***	-0.25	-0.17**	-0.22	0.03	0.04
4	-0.63***	-0.82	-0.49***	-0.64	0.13*	0.17
5	-0.84***	-1.09	-0.62***	-0.81	0.22***	0.29
6	-0.59***	-0.77	-0.37***	-0.48	0.21***	0.27
7	-0.81***	-1.05	-0.53***	-0.69	0.28***	0.36
8	-0.94***	-1.22	-0.91***	-1.18	0.03	0.04
Absolute value of the difference in maximum and minimum value	0.75	0.97	0.74	0.96	0.25	0.32
STAR rating						
1	-0.98***	-1.28	-0.91***	-1.18	0.07	0.09
2	-0.87***	-1.13	-0.72***	-0.94	0.15***	0.20
3	-0.72***	-0.94	-0.55***	-0.72	0.17***	0.22
4	-0.46***	-0.60	-0.26***	-0.34	0.19***	0.25
5	-0.19***	-0.25	-0.20***	-0.26	-0.01	-0.01
Absolute value of the difference in maximum and minimum value	0.79	1.03	0.71	0.92	0.20	0.26
Number of school-year observations	**16 - 224**		**16 - 227**		**16 - 225**	

Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

STAR is School Transparency and Reporting Framework.
* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Note: The table shows the average school-level difference for reports on the 1- to 5-point perseverance scale (described in appendix B) between teachers and students, students and parents, and teachers and parents.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C17. Difference between respondent types overall and by school characteristics for rigorous expectations

Sample	Teacher-student		Teacher-parent		Student-parent	
	Difference in scale	Standardized difference	Difference in scale	Standardized difference	Difference in scale	Standardized difference
All	0.37***	0.47	0.43***	0.55	0.06**	0.08
Ward						
1	0.43***	0.55	0.41***	0.52	-0.02	-0.03
2	0.46***	0.59	0.44***	0.56	-0.03	-0.04
3	0.41***	0.52	0.53***	0.68	0.11***	0.14
4	0.38***	0.49	0.48***	0.61	0.10*	0.13
5	0.43***	0.55	0.53***	0.68	0.10*	0.13
6	0.33***	0.42	0.46***	0.59	0.14***	0.18
7	0.33***	0.42	0.48***	0.61	0.15**	0.19
8	0.27***	0.35	0.26***	0.33	-0.01	-0.01
Absolute value of the difference in maximum and minimum value	0.19	0.24	0.27	0.35	0.18	0.23
STAR rating						
1	0.55***	0.70	0.54***	0.69	0.00	0.00
2	0.36***	0.46	0.44***	0.56	0.08*	0.10
3	0.29***	0.37	0.38***	0.49	0.09*	0.12
4	0.31***	0.40	0.42***	0.54	0.10***	0.13
5	0.43***	0.55	0.49***	0.63	0.07	0.09
Absolute value of the difference in maximum and minimum value	0.26	0.33	0.16	0.20	0.10	0.13
Number of school-year observations	**16 - 224**		**16 - 227**		**16 - 225**	

Absolute value of the standardized difference	Small	Moderate	Substantive	Large
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

STAR is School Transparency and Reporting Framework.
* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Note: The table shows the average school-level difference for reports on the 1- to 5-point rigorous expectations scale (described in appendix B) between teachers and students, students and parents, and teachers and parents.
Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C18. Correlation between school characteristics and average differences in responses for perseverance

Variable	Teacher-student	Teacher-parent	Student-parent
Percentage of English learner students	-0.17**	-0.12	-0.09
Percentage of special education students	0.17*	0.18**	0.17*
Percentage of Black students	0.61***	0.48***	0.21**
Percentage of Hispanic students	-0.22***	-0.14*	-0.08
Percentage of White students	-0.61***	-0.50***	-0.21**
Student survey completion rate	-0.30***	-0.36***	0.01
Number of school-year observations	**223 - 224**	**223 - 224**	**224 - 225**

Absolute value of the correlation	**Low**	**Moderate**	**Substantive**	**High**
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Note: The table shows the pairwise correlation of average absolute school-level differences for perseverance (described in appendix B) between teacher and student reports, student and parent reports, and teacher and parent reports.
Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C19. Correlation between school characteristics and average differences in responses for rigorous expectations

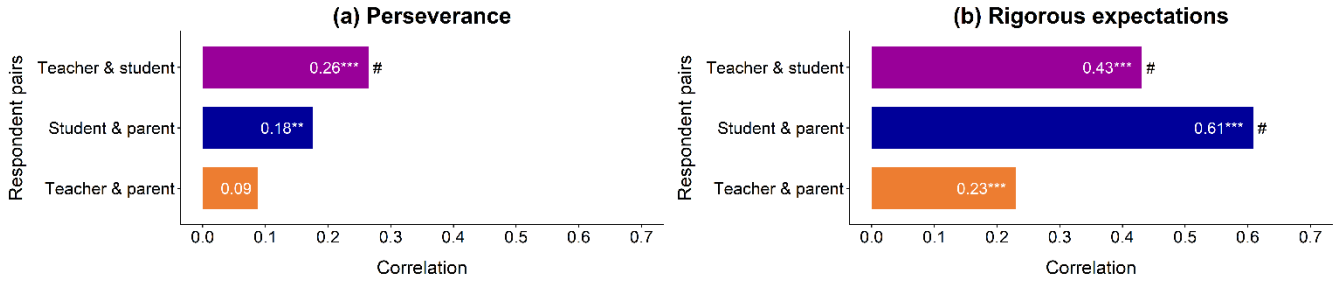
Variable	Teacher-student	Teacher-parent	Student-parent
Percentage of English learner students	-0.04	-0.13	-0.15*
Percentage of special education students	0.05	-0.06	0.39***
Percentage of Black students	0.01	0.04	0.26***
Percentage of Hispanic students	0.00	-0.08	-0.14*
Percentage of White students	-0.02	0.03	-0.22***
Student survey completion rate	-0.43***	-0.25***	-0.25***
Number of school-year observations	**223 - 224**	**223 - 224**	**224 - 225**

Absolute value of the correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

* Significant at $p < .05$; *** significant at $p < .001$.
Note: The table shows the pairwise correlation of average absolute school-level differences for rigorous expectations scale (described in appendix B) between teacher and student reports, student and parent reports, and teacher and parent reports.
Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

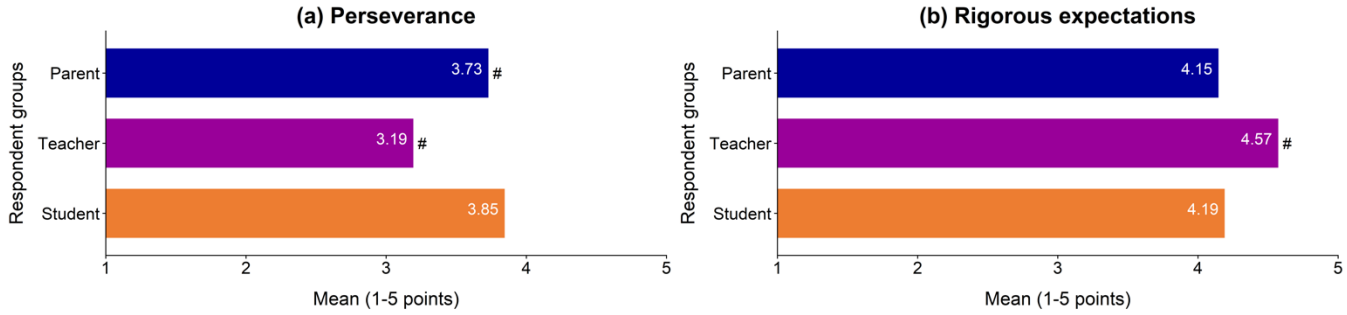
Using nonresponse weights for students and teachers made little difference in the estimated alignment across student, parent, and teacher responses (figures C6 and C7). Relative to the analyses without weights, the correlations between respondent types changed by at most 0.02, and the average differences between respondent types changed by at most 0.01. This finding suggests that student nonresponse bias does not drive the results. However, the study team could not explore teacher and parent nonresponse, so it is not possible to rule out that the findings did not suffer from nonresponse bias due to those respondent types.

Figure C6. Accounting for student and teacher nonresponse made little difference in the estimated correlations between respondent types



** significant at $p < .01$; *** significant at $p < .001$.
# Correlation differs from scale-specific "teacher & parent" correlation by 0.10 or more.
Note: The figure shows the pairwise correlation of school-level averages of perseverance and rigorous expectations scales (described in appendix B) between teacher and student reports, student and parent reports, and teacher and parent reports. The school-level averages for students and teachers were estimated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Figure C7. Accounting for student and teacher nonresponse made little difference in the estimated average differences between respondent types



Absolute value of difference relative to students meets or exceeds 0.10.
Note: The figure shows the average of school-level reports of perseverance and rigorous expectations scales (described in appendix B) for parents, teachers, and students. The school-level averages for students and teachers were estimated using nonresponse weights, as described in appendix B.
Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

References

Almlund, M., Duckworth, A., Heckman, J. J., & Kautz, T. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 1–181). Elsevier.

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.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*(4), 451–464.

Schochet, P. Z. (2008). Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioral Statistics*, *33*(1), 62–87.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*(4), 718–737.

West, M. R., Pier, L., Fricke, H., Loeb, S., Meyer, R. H., & Rice, A. B. (2018). *Trends in student social-emotional learning: Evidence from the CORE districts.* Policy Analysis for California Education.

Appendix D. Additional analyses†

This appendix presents additional analyses for research question 3.

Correlations among changes in SEL competencies, school experiences, and academic measures

To examine the relationships among changes in SEL competencies, school experiences, and academic measures, the study team calculated the correlation between year-to-year changes in each SEL competency and school experience measure and year-to-year changes in the other SEL competencies, school experiences, and key academic measures. DCPS's theory of action highlights that improvements in sense of belonging—a school experience—will eventually lead to improvements in SEL competencies and other school experiences and, in turn, boost students' academic outcomes. To explore this possibility, this analysis estimates correlations between changes in sense of belonging and the other measures, extending research question 3. For completeness, the study team conducted parallel analyses for the other SEL competencies and school experiences. As with the analyses in the main text, the study team used nonresponse weights to account for differences between students who responded to the Panorama survey and those who did not.

Changes in individual SEL competencies and school experiences are highly correlated

The change in each SEL competency and school experience—including sense of belonging—was highly correlated with changes in all other SEL competencies and school experiences, with correlations ranging from 0.31 to 0.51 (table D1). The change in sense of belonging was most highly correlated with changes in the two other measures of school experiences (student satisfaction and rigorous expectations). The results for the other SEL competencies and school experiences exhibited similar patterns to those for sense of belonging.

The correlations among the changes in SEL competencies and school experiences could potentially reflect that—because all SEL competency and school experience items are measured on the same survey—students could have responded to all survey items similarly in a given year. For example, if in the 2018/19 school year a student generally felt positive on the day of the survey, the student might have selected more favorable responses on all items, generating a positive correlation among the changes in the SEL competencies and school experience variables. Evidence suggests that self-reported surveys can suffer from similar kinds of bias. For example, a study of grade 8 students in Boston public schools provided evidence that reference bias contributed to "paradoxical" results, where students in high-performing schools rated their SEL competencies lower (West et al., 2016).

Table D1. Correlations among 2017/18 to 2018/19 changes in SEL competencies and school experiences

		SEL competencies				Other school experiences	
SEL competency/school experience	Sense of belonging	Perseverance	Self-management	Self-efficacy	Social awareness	Rigorous expectations	Student satisfaction
Sense of belonging	--	0.34	0.31	0.36	0.43	0.46	0.51
Perseverance	0.34	--	0.46	0.51	0.47	0.44	0.33
Self-management	0.31	0.46	--	0.49	0.51	0.34	0.32
Self-efficacy	0.36	0.51	0.49	--	0.44	0.44	0.37
Social awareness	0.43	0.47	0.51	0.44	--	0.41	0.40
Rigorous expectations	0.46	0.44	0.34	0.44	0.41	--	0.48
Student satisfaction	0.51	0.33	0.32	0.37	0.40	0.48	--

Absolute value of the correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

SEL is social and emotional learning.
Note: The table shows pairwise correlations between measures in the left column and the measures in the top row. The measures are defined in table B3 in appendix B. The measures were all recoded so that a higher value of the measure is beneficial. The correlations are calculated using nonresponse weights, as described in appendix B. The sample includes students who completed the SEL survey in 2017/18.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Changes in individual SEL competencies and school experiences have low correlations with changes in academic measures

Although the correlations between the change in sense of belonging and changes in academic measures were mostly positive, they were low in all cases (table D2). The results for the other SEL competencies and school experiences exhibited similar patterns to those for sense of belonging.

This evidence suggests that changes in SEL competencies and school experiences are not related to changes in academic measures, but two considerations may limit this interpretation. First, consistent with DCPS's theory of action, changes in school experiences may take time to lead to changes in other variables, especially academic measures. Due to data limitations, this study focuses on correlations between contemporaneous changes in SEL competencies or school experiences and academic measures.[26] If school experiences are more immediately related to students' SEL competencies than academic measures, then changes in sense of belonging may have a stronger correlation with contemporaneous changes in SEL competencies than with those in academic measures. This possibility might arise, for example, if changes in sense of belonging improved students' SEL competencies, and, in turn, changes in SEL competencies affected students' academic measures over time. Second, two of the academic measures—in-seat attendance and days suspended—have limited variation because many students are rarely absent or suspended, which could lead to low correlations due to a restriction of range.

[26] One limitation of this study is that, due to a lack of comparable data, it cannot explore how changes in sense of belonging from 2017/18 to 2018/19 relate to changes in other variables further in the future (for example, changes from 2018/19 to 2019/20). More years of data might help expose this pattern.

Table D2. Correlations among 2017/18 to 2018/19 changes in academic measures

SEL competency/school experience	Academic measures: Math achievement	ELA achievement	In-seat attendance	Days suspended
Sense of belonging	0.02	0.03	0.02	0.04
Perseverance	0.01	0.03	0.04	0.01
Self-management	0.03	0.03	0.02	-0.01
Self-efficacy	0.04	0.02	0.00	0.00
Social awareness	0.03	0.01	0.01	-0.01
Rigorous expectations	0.03	0.02	0.01	0.02
Student satisfaction	0.03	0.01	-0.01	0.01

Absolute value of the correlation	Low	Moderate	Substantive	High
	0.00–0.09	0.10–0.19	0.20–0.29	0.30+

ELA is English language arts. SEL is social and emotional learning.
Note: The table shows pairwise correlations between measures in the left column and the measures in the top row. The measures are defined in table B3 in appendix B. The measures were all recoded so that a higher value of the measure is beneficial. The correlations are calculated using nonresponse weights, as described in appendix B. The sample includes students who completed the SEL survey in 2017/18.
Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Reference

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 students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis, 38*(1), 148–170.