EdWorkingPaper No. 24-900

Are Students On Track?: Comparing the Predictive Validity of Administrative and Survey Measures of Cognitive and Noncognitive Skills for Long-Term Outcomes

Christopher Cleveland
Boston University

Ethan Scherer
Harvard University

Education leaders must identify valid metrics to predict student long-term success. We exploit a unique dataset containing data on cognitive skills, self-regulation, behavior, course performance, and test scores for 8th-grade students. We link these data to data on students' high school outcomes, college enrollment, persistence, and on-time degree completion. Cognitive tests and survey-based self-regulation measures predict high school and college outcomes. However, these relationships become small and lose statistical significance when we control for test scores and a behavioral index. For leaders hoping to identify the best on-track indicators for college completion, the information collected in student longitudinal data systems better predicts both short- and long-run educational outcomes than these survey-based measures of self-regulation and cognitive skills.

VERSION: January 2024

Suggested citation: Cleveland, Christopher, and Ethan Scherer. (2024). Are Students On Track?: Comparing the Predictive Validity of Administrative and Survey Measures of Cognitive and Noncognitive Skills for Long-Term Outcomes. (EdWorkingPaper: 24-900). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/hab2-1m37

Are Students On Track?: Comparing the Predictive Validity of Administrative and Survey

Measures of Cognitive and Noncognitive Skills for Long-Term Outcomes

Christopher Cleveland*
Boston University
chcleve@bu.edu

Ethan Scherer
Center for Education Policy Research at
Harvard University
ethan_scherer@gse.harvard.edu

Abstract

Education leaders must identify valid metrics to predict student long-term success. We exploit a unique dataset containing data on cognitive skills, self-regulation, behavior, course performance, and test scores for 8th-grade students. We link these data to data on students' high school outcomes, college enrollment, persistence, and on-time degree completion. Cognitive tests and survey-based self-regulation measures predict high school and college outcomes. However, these relationships become small and lose statistical significance when we control for test scores and a behavioral index. For leaders hoping to identify the best on-track indicators for college completion, the information collected in student longitudinal data systems better predicts both short- and long-run educational outcomes than these survey-based measures of self-regulation and cognitive skills.

Acknowledgements:

We thank Martin West for project support. We thank the Walton Family Foundation for financial support and the Massachusetts Department of Elementary and Secondary Education and Charter Management Organizations for providing the administrative data for this work. In addition, the research reported here was supported, in whole or in part, by the Institute of Education Sciences, U.S. Department of Education, through grant R305B150010 to Harvard University. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education. Any errors of fact or interpretation are our own.

* Indicates correspondence author

Introduction

Education leaders have traditionally focused on test scores. Research has shown that standardized tests predict later outcomes, but there is a growing belief that they do not provide the whole picture (Kuncel & Hezlett, 2007; Sackett et al., 2008; Goldhaber et al., 2020). Three large-scale studies have documented that high-school GPA dramatically outperforms SAT/ACT admission tests in predicting college graduation, even without adjusting for high-school quality (Geiser & Santelices, 2007; Bowen et al., 2009; Galla et al., 2019). Due to the subjectivity of grades across schools, these results are initially surprising. However, a growing body of research suggests that social and emotional competencies (SEC) (sometimes called non-cognitive competencies), like self-regulation, are differentially crucial for earning course grades compared to standardized tests (Duckworth et al., 2012; Liu et al., 2023). Galla et al. (2019) state that these competencies provide the most incremental predictive power for essential metrics, like on-time graduation. These results help broaden our understanding of what builds future success and have pushed policymakers to explore additional metrics.

The field has proposed several ways to capture these competencies. For example, many districts collect student, teacher, or parent reports of these competencies. However, there are well-documented concerns surrounding self-reported survey measures: (1) social desirability bias because students want to be viewed favorably by their teachers or schools (Duckworth & Yeager, 2015) and (2) outsiders' ratings are generally more predictive of individuals' behaviors than those of the individuals themselves (Connelly & Ones, 2010; Oh et al., 2011; Poropat, 2014) and (3) reference bias – rating is relative to your peers rather than absolute (Duckworth & Yeager, 2015; West et al., 2016). As such, researchers have been exploring other metrics to measure noncognitive competencies. Building on earlier findings of GPA's importance in predicting college graduation, researchers have explored whether school administrative data could be better than self-reports (Kautz & Zanoni, 2014; Heckman et al., 2018; Jackson, 2018). For example, Kautz and Zanoni (2014) create a noncognitive proxy that factors in students' grades, credits, suspension, expulsions, and absences. Liu et al. (2023) use data from CORE and find that academic behaviors, like absences and suspensions, are more predictive than socioemotional competencies for various outcomes, including high school graduation and college enrollment. Others have used participation in extracurricular activities as a proxy for noncognitive skills (Lleras, 2008). However, to our knowledge, no study has computed self-regulation using administrative data and compared its predictive validity for college graduation to the self-reported survey measures.

Relatedly, there is concern that the predictive validity of standardized test scores on post-secondary success may only reflect differences in students' underlying cognitive abilities rather than something more malleable. Cognitive skills like processing speed (how fast to carry out simple cognitive tasks), working memory (the amount of information that can be processed and kept in mind), and fluid reasoning (the flexibility to problem solve in a novel domain) tend to be stable after age 10 (Heckman & Mosso, 2014). If standardized test scores only measure cognitive ability, schools could do little to improve these scores after age 10. However, Finn et al. (2014) find evidence that state standardized tests might measure something different than cognitive skills as they show that schools improved math achievement but did not impact students' cognitive abilities.

Because it is rare to have standardized tests, cognitive skills, and college outcomes in the same dataset, whether test scores predict college success only through cognitive skills remains an undeveloped area of research. We address these unanswered questions by exploiting a unique dataset containing administrative data on behavior and course performance, survey-based measures of self-regulation (i.e., conscientiousness, grit, and self-control), and measures of cognitive skills (both test scores and measures of processing speed, working memory, and fluid reasoning) for 1,338 8th-grade students attending a Northeast district. The sample includes both traditional and charter public schools. We link these data to information from the state longitudinal data system on students' high school outcomes (test performance and on-time graduation) and information from the National Student Clearinghouse (NSC) on college enrollment, persistence, and bachelor's degree completion up to four years after expected high school graduation. Our data lets us understand which measures predict high school and college success.

2. Data and Measures

Administrative Data

We collected administrative data from 8th-grade students attending a Northeast district’s middle schools during the spring semester of the 2010-2011 school year. Within these schools, we sampled all students from whom we received parental consent to participate and who attended school on the data collection day. Students completed cognitive tests and surveys assessing their self-regulation abilities in their regular classrooms.

We then merged these data with student-level high school administrative data—enrollment, attendance, suspensions, grade point average (GPA), math and English language arts (ELA) state test scores, and typical demographic information. The scaled scores were standardized by grade, subject, and year for the whole district to have mean zero and variance one. Of the 3,723 8th-grade students in the district with complete administrative and demographic data, 2,586 attended a surveyed school. We restrict our analytic sample to the 1,338 survey participants with complete administrative and demographic information.

We also received student-level college information from the National Student Clearinghouse (NSC). Our NSC data is for 2012 to 2019, so our college outcomes are restricted to within four years of a student's expected high school graduation. We define college-going as enrolling in a two- or four-year college and college persistence as the number of college quarters attended. A quarter is a student having an enrollment start or end date in one of four three-month periods—January to March, April to June, July to September, and October to December. When defining quarters, enrollment end dates supersede enrollment start dates. We define bachelor's degree completion as a student graduating from a four-year college.

The demographic characteristics of the sample are in Table 1. Table 1 compares demographic and 8th-grade behaviors and outcomes of students between all 8th-grade students to the subset of students who attended a school that administered surveys and to the specific subset of students who consented. Our final analytic sample is racially and ethnically diverse; 71% are Hispanic or Black and come from low-income families, with 81% of students receiving free or reduced-price lunch. These participants differ from the district sample. The surveyed students outscored the other schools by 0.29 standard deviations in math, 0.25 standard deviations in ELA, and 0.19 in GPA. They had fewer absences by 3.22 days and 0.07 fewer suspensions. These differences suggest a positive selection into the survey sample. This selection into survey participation may limit the generalizability of our findings.

Behavioral Index

Our work compares survey and test-based measures with behavioral data already collected regularly by school systems. Following (Kautz & Zanoni, 2014; Heckman et al., 2018;

Jackson, 2018), we proxy for socioemotional competencies by using non-test score behavior available in the state data: the log of the number of absences in 8th grade, whether the student was suspended in 8th-grade and 8th-grade GPA. These behaviors are associated with commonly used self-regulation measures at scale (West et al., 2020). Then, like Jackson (2018), we use a principal component analysis (PCA) to create an index of these measures. We add one to each student's absence count to avoid cases where a student has no absences and the natural log is undefined. For ease of interpretation, we multiply the natural log of absences and suspension indicator by a negative one to combine it with GPA, which should positively affect our outcomes. The PCA retained a single factor (eigenvalue = 1.59). The factor loadings for each of the three variables are as follows: natural log of absences = 0.78, ever suspended = 0.61, and GPA = 0.78. We then predict scores for each student using the Bartlett method and standardize the final index to have a mean zero and variance one (Jackson, 2018).

Measures of Self-Regulation Skills

Our study includes three measures of self-regulation skills: conscientiousness, self-control, and grit.

Conscientiousness

To evaluate conscientiousness, students completed the Big Five Inventory, which assesses five personality traits—neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness (John & Srivastava, 1999). Students rate how much they agree or disagree with a statement (1 = strongly disagree to 5 = strongly agree). Nine items relate to conscientiousness (e.g., "I think I am someone who is a reliable worker"). A student's conscientiousness score is calculated by taking the average of their ratings on these items. This scale had an internal reliability of 0.75.

Self-Control

To assess self-control, students completed the Impulsivity Scale for Children (Tsukayama et al., 2013), which has eight items that measure students' impulsivity as it relates to behavior, attention, and emotions. Items asked students to indicate how often, on a five-point scale ranging from "almost never" to "at least once a day," they exhibited specific behaviors in the past year. Four items assessed interpersonal self-control (e.g., "I interrupted other students while they were talking"), and four assessed intrapersonal self-control (e.g., "I forgot something I needed for class"). We calculated their overall impulsivity score by reverse-coding and averaging these eight items. This scale had an internal reliability of 0.83.

Grit

Grit is "perseverance and passion for long-term goals" (Duckworth et al., 2007) using the Short Grit Scale (Duckworth & Quinn, 2009). Students respond to eight items (e.g., "I finish whatever I begin") on a five-point scale that ranges from "not like me at all" to "very much like me." Overall grit is calculated by taking the average score for these eight items. This scale had an internal reliability of 0.64.

We then use PCA for the three self-regulation skills capturing academic perseverance and refer to the composite of the three skills as our 'self-regulation index.' The PCA retained a single factor (eigenvalue = 2.06). The factor loadings for the three measured self-regulation abilities include conscientiousness = 0.61, impulse control = 0.52, and grit = 0.59. We then predicted scores using the Bartlett method for each student and standardized them to have mean zero and variance one (Jackson, 2018). Item Response Theory-based scores were also developed. The correlation with the PCA is 0.97. Therefore, we present the PCA-based ‘self-regulation index’ in the paper.

Measures of Cognitive Skills

We include three measures of cognitive skills in the study: processing speed, working memory, and fluid reasoning (Finn et al., 2014).

Processing Speed

We use the Coding and Symbol Search subtests from the fourth edition of the Wechsler Intelligence Scale for Children to evaluate processing speed (Wechsler, 2003). For the coding task, students were given a key that assigned the digits 1 through 9 unique symbols and asked to translate a string of digits to the corresponding symbols. For the symbol search task, students were asked to determine whether the two symbols on the left side of the page matched any of the five symbols on the right. Students had two minutes for each task.

Working Memory

We use the count span task to measure working memory (Case, Kurland, & Goldberg, 1982; Cowan et al., 2005). For this task, students were presented with a display of blue circles, blue triangles, and red circles and were instructed to note the number of blue circles. Students viewed as few as one and as many as six such displays before being asked to recall the number of blue circles in each display of the series. The task begins with a single display and increments by one for every three consecutive trials of a given load the student gets correct and maxes out at a load of six. Students were given 4.5 seconds to note the number of blue circles in each display.

Fluid Reasoning

We use the fourth edition of the Test of Nonverbal Intelligence to measure fluid reasoning (Version A; Brown et al., 2010). Students chose which of the six pictures completed the given puzzle. Students were given ten minutes to complete as many as 40 puzzles. Puzzles increased in difficulty as the students progressed through the task.

Due to the high correlation between these three measures, we use a PCA to create a composite of these measures that we refer to as our 'cognitive index.' The PCA retained a single factor (eigenvalue = 1.60). The factor loadings for the three measured cognitive abilities include processing speed = 0.60, working memory = 0.53, and fluid intelligence = 0.60. We then predicted scores using the Bartlett method for each student and standardized them to have mean zero and variance one (Jackson, 2018).

3. Empirical Strategy

To evaluate whether our measures collected in the Spring of 8th grade are predictive of high school and college outcomes, we estimate:

$$Y_{i,t+n} = \beta_1 SR_{it} + \beta_2 C_{it} + \beta_3 BI_{it} + \beta_4 MATH_{it} + \beta_4 ELA_{it} + \delta X_{it} + \sigma_s + \epsilon_{it}$$

$Y_{i,t+n}$ represents student *i*'s outcome in a period after period t. C_{it} representing student i's self-regulation index and cognitive index scores, respectively, at time t in 8th grade, BI_{it} represents the student's behavioral index score, and $MATH_{it}$ and ELA_{it} the student's math and ELA standardized scores. X_i represents a vector of student *i*'s characteristics, including gender, race, ever-economically disadvantaged, ever-special education status, and ever-English language learner. σ_s represents school fixed effects. Standard errors are clustered at the 8th-grade school. Because GPA practices vary across schools and prior research has found that GPA predicts graduation within schools, we ran our models with and without the school fixed effects. We do not find a substantial difference between these two models and present the school fixed effects model.

4. Results

Correlations of Measures

Table 2 shows Pearson correlations among the cognitive measures, self-regulation measures, three composite measures, and 8th-grade ELA and math test scores.

While each cognitive test measures a different aspect of cognitive skills, we see significant correlations (0.27 to 0.36) among them. The Fry and Hale (1996) developmental cascade theory suggests that changes in processing speed are related to changes in working memory that drive changes in fluid intelligence, which makes the weak to moderate relationships we see among these skills fitting. Our three self-regulation measures show moderate to strong correlations (0.42 to 0.67) among the competencies that reflect academic perseverance.

Our behavioral index is most strongly correlated with our self-regulation index (0.30), since the components of the behavioral index—attendance, suspensions, and GPA—largely reflect students' self-regulation competency. When looking at individual self-regulation competencies, the behavioral index correlates most strongly with self-control (0.27) and conscientiousness (0.26). The cognitive index ranks fourth in correlation strength with the behavioral index (0.25), which probably reflects the academic portion measured through GPA.

The 8th-grade test score measures also exhibit significant correlations with other measures. Math scores correlate most with fluid reasoning (0.55) and the overall cognitive index (0.59). ELA scores correlate most with fluid reasoning (0.39) and the overall cognitive index (0.43). The Math and ELA test scores have the lowest correlation with grit, (0.01) and (0.00), respectively.

In sum, cognitive measures correlate strongest with each other, as do self-regulation measures that reflect academic perseverance. The behavioral index correlates strongly with the self-regulation index, providing preliminary evidence that administrative data may be a reasonable proxy for socioemotional competencies. In a separate specification in Appendix Table 1, we exclude GPA from the behavioral index. The behavioral index still correlated most strongly with the non-cognitive index.

High School Outcomes

We first examine how our administrative data and cognitive and self-regulation indices predict 10th-grade math and ELA test scores (Table 3, models 1-5 and 6-10, respectively). For each outcome, the first model only includes administrative data and shows the predictive validity for existing agency data. The next model only includes the cognitive and self-regulation survey measure to demonstrate whether these measures predict our outcomes. In the third model we add the behavioral index with the cognitive and self-reported self-regulation index to assess how the self-reported and the administrative proxy of self-regulation competencies provide independent predictive power of the outcomes, controlling for cognitive skills. Next is a model with 8th-grade test scores, survey measures, and cognitive skills. The final model includes all measures.

In Table 3, models (1), (2), (6), and (7) demonstrate that when either the survey measures or the administrative measures are added to the model, both the cognitive and noncognitive measures independently predict 10th-grade standardized scores in math and ELA. In both cases, the magnitude of the effect for the “non-cognitive” component of the regression is much smaller than the cognitive component but still independently predicts future test score performance. For example, self-control related to attention captured in GPA but not in prior test scores could explain why the behavioral index is independently predictive of test scores even when controlling for prior test scores in models (1) and (6). However, only the administrative construct remains significant when the administrative measures are compared to the survey-based ones. While we would expect the 8th-grade test scores to be strongly correlated with the 10th-grade measure, the fact that the cognitive skills offer no additional information is interesting because test scores are often considered synonymous with cognitive skills. We observe that other measures, like the behavioral index, while small, do provide additional information to the prediction.

We next examine the high school dropout and graduation outcomes (Table 4, models 1-5 and 6-10, respectively). In models (1), (2), (6), and (7), only the noncognitive measures independently predict dropout and graduation. However, when we include both the survey data and the behavioral index, models (3) and (8), respectively, the self-regulation index loses significance, suggesting the administrative data are better proxies for the competencies than the survey measures.

In model (9), we remove and replace the behavioral index with test scores. Math test scores, but not ELA test scores, explain additional variance in high school graduation but not dropout. However, passing the 10th-grade standardized test is required to graduate high school in the state. However, this requirement is for both Math and ELA test scores, and the relationship between ELA scores and graduation is insignificant. Thus, the Math test scores could be proxied for cognitive and noncognitive skills other than self-regulation because our self-regulation index remains statistically significant and unchanged. However, as we observed in regression (6), there is no independent predictive power of math test scores once the behavioral index is included. Graduating from high school involves getting homework in on time and attending classes, which requires self-regulation, among other competencies. Thus, its dominance over all the other measures seems reasonable.

College Outcomes

We next examine two- or four-year college enrollment, college persistence, and bachelor's degree completion (Table 5, models 1-5 6-10, and 11-15, respectively). In models (1), (2), (6), and (7), we observe that the noncognitive competencies most strongly predict college enrollment and persistence, which likely requires students to submit the appropriate forms on time, attend classes, and manage their newfound freedom or constraints among other logistical hurdles. What is surprising is that 8th grade ELA is independently predictive of college enrollment and persistence. College essays and enrollment require correspondence, and most majors require some writing; however, why ELA is predictive but not mathematics is unexpected. Further, as we observed with graduation and dropout, once the behavioral index is added to the regression, the self-regulation survey index is no longer independently predictive for enrollment, but this is not the case for persistence in college. These findings differ from what Liu et al. (2023) found, though our exact survey measures differ, and the study occurred in a different country region.

The behavioral index and math are significant for bachelor's degree completion, but not ELA, in model (11). However, examining the magnitude of the ELA effects, they are similar but slightly smaller than math. When we exclude the behavioral index and test scores, the cognitive and self-regulation indices are significant predictors in model (12). When we have all the predictors in the model, the behavioral index and math scores are the only significant predictors (15). The fact that GPA, absence, and suspension occurring eight years prior can predict a nine-percentage point increase in on-time bachelor's degrees demonstrates the power of the tools that schools already have in their hands. In a separate specification in which the behavioral index does not include GPA, the behavioral index coefficient diminishes by approximately 50 percent and is only significant at the 0.10 level. However, GPA retains a strong relationship, yet for enrollment and persistence, the statistical significance remains. This change highlights the added predictive power of absences, suspensions, and GPA (see Appendix Table 2). We also separately examine college persistence only among the students who enrolled in college. Here, the behavioral index is the only covariate that predicts college persistence, but for on-time 4-year college degree completion, 8th grade math test scores remain independently predictive. Thus, the self-regulation skills proxied by these behavioral data remain predictive even among college enrollees (see Appendix Table 3).

In all models for all outcomes, survey measures alone explain the slightest variation and have the worst fit based upon adjusted R^2, Akaike information criterion (AIC), and Bayesian information criterion (BIC). Other than math, models with administrative data alone explain the most variation and have the best fit.

5. Discussion and Conclusion

Over the last two decades, evidence has mounted that success in post-secondary education and beyond relies on cognitive and noncognitive skills not fully captured by standardized tests. The increasing evidence of the importance of noncognitive skills, like self-regulation, has pushed policymakers to collect and use student self-report surveys. The collection of these data has the potential to assess whether some students are on track to enroll in college and complete a degree. However, recent work has also questioned whether these survey-based measures are the best metric to assess student preparation or if behavioral proxies of noncognitive skills already collected through administrative datasets better predict these outcomes of interest (Liu et al, 2023). Evidence comparing these measures is significant as policymakers continue refining and identifying the skills contributing to college success.

Our work builds on the evidence that cognitive and noncognitive skills measures among diverse student populations predict future educational attainment. We replicate research that finds noncognitive measures tend to outperform cognitive assessments in predicting degree completion (Geiser & Santelices, 2007; Bowen et al., 2009; Jackson, 2018; Galla et al., 2019). A standard deviation increase in the behavioral index is associated with a nine-percentage point increase in the likelihood of obtaining a bachelor's degree. In contrast, a similar increase in 8th-grade mathematics test scores is less than half that amount, and ELA test scores are not independently predictive of degree completion. Thus, developing both skills, particularly noncognitive competencies, predicts post-secondary success.

We extend the literature by assessing measures before high school entry and find that they predict post-secondary enrollment, persistence, and completion. Prior work had either focused on elementary measures to predict high school outcomes (Goldhaber et al., 2020); high school GPA as a predictor of high school graduation (Easton et al., 2007); college-going and retention (Geiser & Stanelices, 2007; Easton et al., 2007; Liu et al, 2023); or tests taken after secondary school to predict further education and earnings outcomes (Kuncel & Hezlett, 2007; Sackett et al., 2008). Our assessment spans early secondary through a bachelor's degree and finds that measures collected early can be highly predictive of high school test score performance, dropout and graduation, and post-secondary outcomes.

Furthermore, by comparing the administrative and survey-based measures in the cognitive and self-regulation domains, we can assess which measures are most predictive for different outcomes. For both high school and college outcomes, once we add the 8th-grade behavioral index and test scores to our prediction regressions with the cognitive and self-reported self-regulation indices, the self-regulation measures lose their statistical significance, suggesting they do not add explanatory power over what is available in the 8th-grade administrative data. Because our dataset contains these four measures, we can also begin to unpack what competencies are captured in administrative data. Several recent studies have demonstrated the association between self-regulation and grade point average (Duckworth et al., 2012; Galla et al., 2019, West et al., 2020). We observe a similar pattern for our behavioral index, combining absences, suspensions, and 8th-grade grade point average. However, the index also strongly correlates with our cognitive index, combining working memory, processing speed, and fluid reasoning. Thus, while an index with administrative measures such as absences, suspensions, and grade point average appear to proxy for self-regulation skills, the correlation of our cognitive index is almost as strong as the self-regulation index.

We also find evidence that absences and suspensions remain predictive of long-run outcomes even without GPA in our behavioral index. Surprisingly, when we remove GPA from our behavior index and use it as a separate covariate in supplemental analyses, it remains predictive of post-secondary enrollment and persistence. Thus, even though most prior work has focused on GPA as a predictor of college success, our evidence indicates that adding other administrative measures like absences and suspensions can strengthen this measure because they independently predict post-secondary outcomes (Geiser & Santelices, 2007; Bowen et al., 2009; Galla et al., 2019). We unpack this further when we explore the correlation of the modified behavioral index with our survey measures. It correlates most strongly with impulse control, whereas the index including GPA correlates as strongly with impulse control as it does grit and conscientiousness. Ultimately, the administrative behavioral index captures many aspects of student skills independent of test scores, perhaps why it is a strong predictor of post-graduate success.

Finally, we find evidence that test scores as predictors of post-secondary success operate through more than cognitive skills alone. In regressions where we control for the self-regulation index, cognitive index, and 8th-grade test scores, the cognitive index is no longer significant, suggesting it does not add explanatory power over what is available in the 8th test scores.

Standardized tests are not only measures of a student's aptitude. The tests capture a student's accumulated knowledge based on state content standards in core academic subjects from educational experiences, teaching, and aptitude. Furthermore, there is evidence that metadata captured by standardized tests can also capture noncognitive skills like self-efficacy, self-regulation, conscientiousness, and grit (Soland et al., 2019). We observe evidence of this finding in our data. While standardized tests are correlated most strongly with processing speed and fluid reasoning, they correlate with impulse control. The fact that test scores are independently predictive of post-secondary success, even controlling for the cognitive index, is comforting, given prior research that schools might have limited ability to improve cognitive skills but might be able to improve standardized tests (Finn et al., 2014).

Policy Implications

Our findings have several policy implications. Our main contribution is to help policymakers and educators best assess which measures predict on-time degree completion.

First, administrative data can provide valuable early warning indicators for students. Administrative data collected in 8th-grade predicts both high school and college outcomes. These measures could target interventions for students. Furthermore, the magnitude of the effects between test scores and the behavioral index differs in expected ways. The behavioral index proxies for both cognitive and noncognitive skills, while test scores proxies for cognitive skills provide information on both developmental constructs. These results demonstrate that practitioners have a powerful tool early in a student's secondary education that predicts short- and long-term outcomes. These measures could help practitioners monitor progress toward high school and college success annually.

Second, in a race between the cognitive and self-regulation survey-based measures and administrative data, the administrative data explain more of the variation in our college outcomes. Particularly for the policymakers hoping to identify the best on-track indicators, the information collected in their current student longitudinal systems better predicts both short- and long-run educational outcomes. Other recent work has emphasized the importance of precisely leveraging the state longitudinal systems for this purpose (Austin et al., 2020; Goldhaber et al., 2020). These results also help researchers and policymakers because they do not need to rely upon relatively new measures, like self-assessed self-regulation, to evaluate prior interventions, assess school progress, or create prediction models for postsecondary enrollment and completion. For agencies, both time and money are scarce resources; thus, diverting staff time to compile and better utilize existing data could be beneficial. Furthermore, the increased focus on noncognitive skills in school systems over the last decade has been because they predict academic outcomes and because, unlike traditional measures of students' success (like standardized tests and cognitive skill), noncognitive skills appear to be more malleable (Heckman & Mosso, 2014; Dee & West, 2011). Thus, agencies could capitalize on existing data to monitor and evaluate interventions, hoping to improve these skills.

Still, test scores and our behavioral index are blunt instruments for understanding the change mechanism. For example, consider if an agency decides to implement a new curriculum to improve teacher and student relationships, which results in a decline in absences. It is challenging to know how to improve the program or scale it effectively without additional data. Similarly, past research has noted that test scores are difficult to change. Thus, while our data demonstrate the superiority of administrative data over self-reported measures in predicting high school and college outcomes, we caution against abandoning survey-based measures entirely.

Future Research

Our work needs further validation in future work. First, we only test our hypotheses on a subset of potential cognitive and noncognitive survey measures. Our measures or related measures are standard, but RAND identified almost 200 noncognitive measures in their literature scan on noncognitive skills (Hamilton et al., 2018). Other metrics should be assessed to see if they yield similar patterns. Second, while we can draw on rich cognitive and noncognitive data, our sample of students performs better than those in the same schools who did not answer the survey items or take the cognitive tests. Thus, further work on these harder-to-reach students could be beneficial. Finally, these findings should be replicated with other populations as recent work highlights that these trends will likely extend to other states and populations (Austin et al., 2020; Goldhaber et al., 2020; Jiu et al., 2023). The replication of our findings in other cities could be beneficial.

References

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. (2011) Personality psychology and economics. In: Hanushek EA, Machin S, Woessmann L, eds. Handbook of the Economics of Education. Amsterdam, Netherlands: Elsevier; 2011:1–181.

Austin, Wes, David Figlio, Dan Goldhaber, Eric Hanushek, Tara Kilbride, Cory Koedel, Jaeseok Sean Lee, Jin Lou, Umut Özek, Eric Parsons, Steven Rivkin, Tim Sass, and Katharine Strunk (2020). Where are Initially Low-performing Students the Most Likely to Succeed? A Multi-state Analysis of Academic Mobility (Preliminary Draft). CALDER Working Paper No. 227-0220.

Bowen, William G, Matthew M. Chingos, and Michael S. McPherson (2009). Crossing the finish line: Completing college at America's public universities. Princeton University Press.

Brown, Linda, Rita J. Sherbenou, and Susan K. Johnsen, (2010). Test of Nonverbal Intelligence—Fourth edition (Version A). San Antonio, TX: The Psychological Corp.

Case, Robbie, D. Midian Kurland, and Jill Goldberg (1982). Operational efficiency and the growth of short-term memory span. Journal of Experimental Child Psychology, 33, 386–404.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. The Quarterly Journal of Economics. 126(4), 1593-1660.

Connelly, Brian S., and Deniz S. Ones (2010). Another perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. Psychological Bulletin, 136(6), 1092-1122.

Dee, Thomas S., and Martin R. West (2011). The noncognitive returns to class size. Education Evaluation and Policy Analysis, 33(1), 23-46.

Duckworth, Angela Lee, and Patrick D. Quinn (2009). Development and validation of the Short Grit Scale (Grit-S). Journal of Personality Assessment, 91, 166–174.

Duckworth, Angela Lee, Christopher Peterson, Michael D. Matthews, and Dennis R. Kelly (2007). Grit: Perseverance and passion for long-term goals. Journal of Personality and Social Psychology, 92, 1087–1101.

Duckworth, Angela Lee, Patrick D. Quinn, and Eli Tsukayama (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. Journal of educational psychology, 104(2), 439.

Duckworth, Angela Lee, and David S. Yeager (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. Educational Researcher, 44(4), 237-251.

Dweck, Carol (2006). Mindset: The new psychology of success. New York, NY: Random House.

Easton, John Q., Esperanza Johnson, and Lauren Sartain (2017). The predictive power of ninth-grade GPA. Chicago, IL: University of Chicago Consortium on School Research.

Finn, Amy S., Matthew A. Kraft, Martin R. West, Julia A. Leonard, Crystal E. Bish, Rebecca E. Martin, Margaret A. Sheridan, Christopher F. O. Gabrieli, and John D. E. Gabrieli (2014). Cognitive Skills, Student Achievement Tests, and Schools. Psychological Science, 25(3), 736–744. https://doi.org/10.1038/jid.2014.371

Fry, Astrid F., and Sandra Hale (1996). Processing speed, working memory, and fluid intelligence: Evidence for a Developmental Cascade. Psychological Science, 7, 237-241.

Galla, Brian. M., Elizabeth P. Shulman, Benjamin D. Plummer, Margo Gardner, Stephen J. Hutt, J. Parker Goyer, Sidney K. D'Mello, Amy S. Finn, and Angela L. Duckworth (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 Education Research Journal, 20(10), 1-39.

Geiser, Saul, and Maria V. Santelices (2007). Validity of High-School Grades in Predicting Student Success beyond the Freshman Year: High-School Record vs. Standardized Tests as Indicators of Four-Year College Outcomes. Research & Occasional Paper Series: CSHE. 6.07. Center for studies in higher education.

Goldhaber, Dan, Malcolm Wolff, and Timothy Daly (2020). Assessing the Accuracy of Elementary School Test Scores as Predictors of Students' High School Outcomes. CALDER Working Paper No. 235-0520

Heckman, James J., Eric Humphries, and Gregory Veramendi (2018). Returns to education: The causal effects of education on earnings, health, and smoking. Journal of Political Economy, 126(S1), S197-S246.

Heckman, James J., and Stefano Mosso (2014). The economics of human development and social mobility. Annu. Rev. Econ., 6(1), 689-733.

Jackson, Kirabo (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes. Journal of Political Economy, 126, 2072-2107.

John, Oliver P., and Sanja Srivastava (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), Handbook of personality: Theory and research (2nd ed., pp. 102–138). New York, NY: Guilford Press.

Kautz, T. D. & Zanoni, W. (2014). Measuring and fostering noncognitive skills in adolescence: Evidence from Chicago Public Schools and the OneGoal Program. Unpublished manuscript, Department of Economics, University of Chicago, Chicago, IL.

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. Science.

Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. Social Science Research 37 (3), 888–902.

Liu, J., Kuhfeld, M., & Lee, M. (2023). Noncognitive Factors and Student Long-Run Success: Comparing the Predictive Validity of Observable Academic Behaviors and Social-Emotional Skills. Educational Policy, 0(0). https://doi.org/10.1177/08959048231209262

Oh, I. S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. Journal of Applied Psychology, 96(4), 762-773.

Poropat, A. E. (2014). Other-rated personality and academic performance: Evidence and implication. Learning and Individual Differences, 34, 24-32.

Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. American Psychologist, 63(4), 215.

Soland, J., Zamarro, G., Cheng, A., & Hitt, C. (2019). Identifying naturally occurring direct assessments of social-emotional competencies: The promise and limitations of survey and assessment disengagement metadata. Educational Researcher, 48(7), 466-478.

Tsukayama, E., Duckworth, A. L., & Kim, B. (2013). Domain-specific impulsivity in school-age children. Developmental Science, 16, 879–893.

Wechsler, D. (2003). Wechsler Intelligence Scale for Children—Fourth edition. San Antonio, TX: The Psychological Corp.

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' noncognitive skills and the impact of schooling. Education Evaluation and Policy Analysis, 38(1), 148-170.

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.

Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. Journal of Educational Measurement, 39(1), 1-37.

Tables & Figures

Table 1: Mean Demographic Characteristics and Academic Indicators for 8th-grade Students Among All Public Schools in the District, Sampled Schools, and Sampled Students

	(1)	(2)	(3)
	All Students in the District	All Students in Sampled Schools	Sampled Students
Male	0.51	0.50	0.45
African American	0.37	0.35	0.35
Asian/Pacific Islander	0.10	0.09	0.10
Hispanic	0.37	0.38	0.36
White, Non-Hispanic	0.15	0.16	0.18
Other	0.02	0.02	0.01
Free/reduced price lunch	0.84	0.83	0.81
Individualized education plan	0.22	0.21	0.18
English language learner	0.18	0.17	0.15
Absences	11.13	10.46	7.91
8th-grade suspensions	0.24	0.23	0.17
8th-grade GPA (standardized)	0.00	0.00	0.19
8th-grade Math scores (standardized)	0.00	0.10	0.29
8th-grade ELA scores (standardized)	0.00	0.09	0.25
Number of students	3723	2586	1338

Notes: All samples are restricted to students with complete demographic and academic information. Sampled schools participate in the cognitive and noncognitive surveys; sampled students have valid data on the cognitive and noncognitive subtests. 8th-grade GPA is standardized within school and across all District students. Math and ELA test scores are standardized across all 8th grade students in the District in 2011 to have mean zero and variance one. Eighth grade suspension is a dummy whether the student was ever suspended that year. Absences is the number of absences. GPA = grade point average. ELA = English language arts.

Table 2. Correlation Matrix of 8th-grade Measures Used in Predictive Models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	PS	WM	FR	Grit	SC	C	CI	SRI	BI	8th Gr Math	8th Gr ELA
Processing Speed (PS)	1										
Working Memory (WM)	0.27**	1									
Fluid Reasoning (FR)	0.36**	0.26**	1								
Grit	-0.03	0.03	-0.04	1							
Self-control (SC)	0.03	0.07*	0.12**	0.42**	1						
Conscientiousness (C)	0.03	0.04	-0.02	0.67**	0.49**	1					
Cognitive Index (CI)	0.77**	0.67**	0.75**	-0.02	0.10**	0.02	1				
Self-Reg Index (SRI)	0.02	0.05	0.02	0.85**	0.74**	0.88**	0.04	1			
Behavioral Index (BI)	0.22**	0.14**	0.19**	0.22**	0.27**	0.26**	0.25**	0.30**	1		
8th-grade Math	0.46**	0.26**	0.55**	0.01	0.12**	0.05	0.59**	0.07*	0.39**	1	
8th-grade ELA	0.37**	0.16**	0.39**	0.00	0.07**	0.03	0.43**	0.04	0.30**	0.70**	1

Notes: All individual cognitive and noncognitive subtest scores are standardized to have mean zero and unit variance. The cognitive index is a weighted average of the three cognitive subtest scores calculated by running a principal components analysis on the three subtests. Similarly, the noncognitive index is a weighted average of students' grit, conscientiousness, and self-control subtest scores and is calculated by running a principal components analysis on these three subtests. The behavioral index is calculated by running a principal components analysis using the natural log of absences, an indicator for whether a student was suspended in 8th-grade, and standardized GPA, where GPA is standardized within school and across all 8th grade students in the District. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Predicting High School Test Scores Using Cognitive and Self-Regulation Survey Scores, Academic, and Behavioral Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	10th-Grade Math Scores					10th-Grade ELA Scores				
Self-Reg Index		0.05**	0.01	0.00	-0.01		0.05**	0.02	0.02	0.01
		(0.012)	(0.014)	(0.012)	(0.012)		(0.012)	(0.014)	(0.009)	(0.010)
Cognitive Index		0.22**	0.18**	0.04	0.04		0.12**	0.09**	0.00	0.00
		(0.046)	(0.042)	(0.019)	(0.019)		(0.028)	(0.026)	(0.017)	(0.017)
Behavioral Index	0.06*		0.26**		0.06**	0.07*		0.22**		0.07*
	(0.024)		(0.037)		(0.023)	(0.029)		(0.035)		(0.031)
8th-grade Math	0.57**			0.57**	0.55**	0.13**			0.15**	0.13**
	(0.053)			(0.049)	(0.047)	(0.020)			(0.018)	(0.019)
8th-grade ELA	0.19**			0.20**	0.19**	0.57**			0.58**	0.57**
	(0.024)			(0.024)	(0.024)	(0.036)			(0.039)	(0.036)
Observations	1060	1060	1060	1060	1060	1069	1069	1069	1069	1069
Adjusted R-squared	0.75	0.56	0.60	0.75	0.76	0.69	0.50	0.53	0.68	0.69
AIC	1388	2015	1916	1393	1386	1660	2164	2103	1671	1664
BIC	1448	2070	1976	1458	1455	1720	2218	2163	1735	1733

Note: Standard errors are reported in parentheses and clustered at the 8th-grade middle school attended. Samples are restricted to students with complete demographic, cognitive, and noncognitive surveys, 8th-grade behavioral and academic outcome data, and a non-missing outcome. All regressions control for student gender and race/ethnicity and for whether a student ever received free- or reduced-price lunch, ever had an individualized education plan, and ever was an English language learner. All models also include school fixed effects. All survey measures are standardized within the sample of students who completed the survey. We standardize the test scores across all students in given grade in the District. *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$.

Table 4: Predicting High School Dropout and Graduation Using Cognitive and Self-Regulation Survey Scores, Academic, and Behavioral Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	High School Dropout					On-Time High School Graduation				
Self-Reg Index		-0.02**	-0.01	-0.02**	-0.01		0.04***	0.01	0.03**	0.01
		(0.007)	(0.006)	(0.007)	(0.007)		(0.009)	(0.007)	(0.009)	(0.008)
Cognitive Index		-0.00	0.01	0.01	0.01		0.01	-0.02	-0.02	-0.02*
		(0.006)	(0.005)		(0.006)		(0.010)	(0.009)	(0.010)	(0.010)
Behavioral Index	-0.08***		-0.08**		-0.08**	0.16***		0.16***		0.16***
	(0.021)		(0.022)		(0.021)	(0.027)		(0.027)		(0.027)
8th-grade Math	0.01			-0.02	0.01	-0.01			0.07**	0.00
	(0.012)			(0.014)	(0.013)	(0.019)			(0.020)	(0.020)
8th-grade ELA	-0.01			-0.02	-0.01	0.02			0.04	0.02
	(0.015)			(0.015)	(0.015)	(0.020)			(0.020)	(0.020)
Observations	1208	1208	1208	1208	1208	1208	1208	1208	1208	1208
Adjusted R-squared	0.109	0.070	0.112	0.076	0.111	0.195	0.101	0.197	0.118	0.197
AIC	320.05	370.83	316.64	365.66	319.73	1009.19	1141.46	1005.59	1119.61	1008.17
BIC	381.21	426.89	377.80	431.92	391.09	1070.35	1197.53	1066.76	1185.87	1079.53

Note: Standard errors are reported in parentheses and clustered at the 8th-grade middle school attended. Samples are restricted to students with complete demographic, cognitive, and noncognitive surveys, 8th-grade behavioral and academic outcome data, and a non-missing outcome. All regressions control for student gender and race/ethnicity and for whether a student ever received free- or reduced-price lunch, ever had an individualized education plan, and ever was an English language learner. All models also include school fixed effects. All survey measures are standardized within the sample of students who completed the survey. We standardize the test scores across all 8th grade students in the District. *** p< 0.01, ** p< 0.05, * p< 0.1

Table 5. Predicting College Outcomes Using Cognitive and Self-Regulation Survey Scores, Academic, and Behavioral Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	College Enrollment					Quarters					Bachelor's degree				
Self-Reg Index		0.03*	0.01	0.02	0.01		0.74***	0.27*	0.60***	0.26*		0.03***	0.01	0.03**	0.01
		(0.01)	(0.01)	(0.01)	(0.01)		(0.16)	(0.11)	(0.15)	(0.11)		(0.01)	(0.01)	(0.01)	(0.01)
Cognitive Index		0.01	-0.01	-0.01	-0.02		0.16	-0.20	-0.28	-0.33		0.02*	0.01	-0.00	-0.00
		(0.01)	(0.01)	(0.01)	(0.01)		(0.19)	(0.16)	(0.21)	(0.19)		(0.01)	(0.01)	(0.01)	(0.01)
Behavioral Index	0.11***		0.13***		0.11***	2.47***		2.59***		2.34***	0.10***		0.11***		0.09***
	(0.02)		(0.02)		(0.02)	(0.31)		(0.30)		(0.28)	(0.02)		(0.02)		(0.02)
8th-Grade Math	0.02			0.07**	0.02	0.01			1.12**	0.20	0.04*			0.08***	0.04*
	(0.02)			(0.02)	(0.02)	(0.27)			(0.31)	(0.30)	(0.02)			(0.01)	(0.02)
8th-Grade ELA	0.05*			0.07**	0.05*	0.93*			1.24**	0.92*	0.03			0.04	0.03
	(0.02)			(0.02)	(0.02)	(0.41)			(0.41)	(0.40)	(0.02)			(0.02)	(0.02)
Observations	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338
Adjusted R-squared	0.208	0.150	0.202	0.176	0.209	0.293	0.222	0.290	0.247	0.296	0.319	0.276	0.314	0.298	0.320
AIC	1276.15	1370.32	1287.17	1330.72	1277.20	8775.29	8903.75	8780.61	8860.27	8772.14	1031.65	1112.77	1042.25	1074.16	1032.54
BIC	1338.54	1427.51	1349.56	1398.31	1349.99	8837.68	8960.94	8843.00	8927.86	8844.93	1094.04	1169.96	1104.63	1141.75	1105.33

Notes: Standard errors are reported in parentheses and clustered at the 8th-grade middle school attended. Samples are restricted to students with complete demographic, cognitive and noncognitive survey, 8th-grade behavioral and academic outcome data, and a non-missing outcome. All regressions control for student gender and race/ethnicity and for whether a student ever received free- or reduced-price lunch, ever had an individualized education plan, and ever was an English language learner. All models also include school fixed effects. All survey measures are standardized within the sample of students who completed the survey. We standardize the test scores across all 8th grade students in the District. *** p< 0.01, ** p< 0.05, * p< 0.1,

Appendix

Appendix Table 1. Correlation Matrix of 8th-grade Measures Where the Behavioral Index Excludes GPA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variable	PS	WM	FR	Grit	SC	C	CI	SRI	SAI	GPA	8th Gr Ma	8th Gr ELA
Processing Speed (PS)	1											
Working Memory (WM)	0.27***	1										
Fluid Reasoning (FR)	0.36***	0.26***	1									
Grit	-0.03	0.03	-0.04	1								
Self-control (SC)	0.03	0.07*	0.12***	0.42***	1							
Conscientiousness (C)	0.03	0.04	-0.02	0.67***	0.49***	1						
Cognitive Index (CI)	0.77***	0.67***	0.75***	-0.02	0.10***	0.02	1					
Self-Reg Index (SRI)	0.02	0.05	0.02	0.85***	0.74***	0.88***	0.04	1				
Susp-Abs Index (SAI)	0.13***	0.10***	0.15***	0.08**	0.17***	0.08**	0.18***	0.13***	1			
GPA	0.22***	0.13***	0.15***	0.30***	0.30***	0.36***	0.23***	0.38***	0.33***	1		
8th-grade Math	0.46***	0.26***	0.55***	0.01	0.12***	0.05	0.59***	0.07*	0.26***	0.35***	1	
8th-grade ELA	0.37***	0.16***	0.39***	0.00	0.07**	0.03	0.43***	0.04	0.17***	0.30***	0.70***	1

Notes: All individual cognitive and noncognitive subtest scores are standardized to have mean zero and unit variance. The cognitive index is a weighted average of the three cognitive subtest scores calculated by running a principal components analysis on the three subtests. Similarly, the noncognitive index is a weighted average of students' grit, conscientiousness, and self-control subtest scores and is calculated by running a principal components analysis on these three subtests. The behavioral index is calculated by running a principal components analysis using the natural log of absences, an indicator for whether a student was suspended in 8th-grade, and standardized GPA, where GPA is standardized within school and across all 8th grade students in the District.
*** p< 0.01, ** p< 0.05, * p< 0.1

Appendix Table 2. Predicting College Outcomes Using Cognitive and Self-Regulation Survey Scores, Academic, and Behavioral Data & Behavioral Index Excludes GPA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Enrolled					Quarters					Bachelor's degree				
Self-Reg		0.03*	0.00	0.02	0.00		0.74***	0.18	0.60***	0.20		0.03***	0.01	0.03**	0.01
Index		(0.01)	(0.01)	(0.01)	(0.01)		(0.16)	(0.11)	(0.15)	(0.11)		(0.01)	(0.01)	(0.01)	(0.01)
Cognitive		0.01	-0.01	-0.01	-0.02		0.16	-0.26	-0.28	-0.33		0.02*	0.00	-0.00	-0.00
Index		(0.01)	(0.01)	(0.01)	(0.01)		(0.19)	(0.17)	(0.21)	(0.19)		(0.01)	(0.01)	(0.01)	(0.01)
Behavioral	0.05***		0.05***		0.05***	1.10***		1.07***		1.10***	0.03		0.03		0.03
Index w/o GPA	(0.01)		(0.01)		(0.01)	(0.17)		(0.17)		(0.17)	(0.01)		(0.01)		(0.01)
8th grade	0.09**		0.10***		0.08***	1.84***		1.98***		1.71***	0.09***		0.10***		0.09***
GPA	(0.02)		(0.02)		(0.02)	(0.30)		(0.33)		(0.29)	(0.02)		(0.01)		(0.02)
8th-Grade	0.01			0.07**	0.02	-0.16			1.12**	0.06	0.03			0.08***	0.03
Math	(0.02)			(0.02)	(0.02)	(0.29)			(0.31)	(0.32)	(0.02)			(0.01)	(0.02)
8th-Grade	0.05*			0.07**	0.05*	0.81			1.24**	0.82*	0.02			0.04	0.02
ELA	(0.02)			(0.02)	(0.02)	(0.40)			(0.41)	(0.40)	(0.02)			(0.02)	(0.02)
Observations	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338	1338
Adjusted R-squared	0.210	0.150	0.206	0.176	0.210	0.292	0.222	0.291	0.247	0.294	0.322	0.276	0.319	0.298	0.321
AIC	1274.69	1370.32	1281.33	1330.72	1276.42	8777.77	8903.75	8780.53	8860.27	8776.41	1027.02	1112.77	1032.26	1074.16	1029.90
BIC	1342.28	1427.51	1348.92	1398.31	1354.40	8845.36	8960.94	8848.11	8927.86	8854.39	1094.60	1169.96	1099.85	1141.75	1107.88

Notes: Standard errors are reported in parentheses and clustered at the 8th-grade middle school attended. Samples are restricted to students with complete demographic, cognitive, and noncognitive surveys, 8th-grade behavioral and academic outcome data, and a non-missing outcome. All regressions control for student gender and race/ethnicity and for whether a student ever received free- or reduced-price lunch, ever had an individualized education plan, and ever was an English language learner. All models also include school fixed effects. All survey measures are standardized within the sample of students who completed the survey. We standardize the test scores across all 8th grade students in the District. *** p< 0.01, ** p< 0.05, * p< 0.1,

Appendix Table 3. Predicting College Outcomes Using Cognitive and Self-Regulation Survey Scores, Academic, and Behavioral Data for College Enrollees

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Quarters					Bachelor's degree				
Self-Reg Index		0.55***	0.25*	0.49***	0.24*		0.04***	0.02	0.03**	0.02
		(0.11)	(0.10)	(0.11)	(0.11)		(0.01)	(0.01)	(0.01)	(0.01)
Cognitive Index		0.08	-0.14	-0.13	-0.16		0.03*	0.01	0.00	-0.00
		(0.13)	(0.12)	(0.14)	(0.13)		(0.01)	(0.01)	(0.01)	(0.01)
Behavioral Index	1.86***		1.80***		1.74***	0.11***		0.12***		0.10***
	(0.37)		(0.38)		(0.37)	(0.02)		(0.02)		(0.02)
8th-Grade Math	-0.23			0.52	-0.13	0.05*			0.09***	0.05*
	(0.30)			(0.26)	(0.29)	(0.02)			(0.02)	(0.02)
8th-Grade ELA	0.58			0.73	0.56	0.03			0.04	0.03
	(0.48)			(0.49)	(0.48)	(0.03)			(0.03)	(0.03)
Observations	985	985	985	985	985	985	985	985	985	985
Adjusted R-squared	0.182	0.138	0.182	0.147	0.183	0.294	0.253	0.289	0.274	0.295
AIC	6187.96	6238.59	6187.04	6229.82	6188.01	966.12	1020.84	973.40	995.17	967.14
BIC	6246.68	6292.41	6245.75	6293.43	6256.51	1024.83	1074.65	1032.11	1058.78	1035.64

Notes: Standard errors are reported in parentheses and clustered at the 8th-grade middle school attended. Samples are restricted to students with complete demographic, cognitive, and noncognitive surveys, 8th-grade behavioral and academic outcome data, and a non-missing outcome. All regressions control for student gender and race/ethnicity and for whether a student ever received free- or reduced-price lunch, ever had an individualized education plan, and ever was an English language learner. All models also include school fixed effects. All survey measures are standardized within the sample of students who completed the survey. We standardize the test scores across all 8th grade students in the District. *** p< 0.01, ** p< 0.05, * p< 0.1,