Does An Early Social-Emotional Learning Program Have Lasting Academic Effects Through High School?

High School Follow-Up Findings from a Randomized Trial of INSIGHTS

JUNE 2026

Pei Zhu
Peyton Nash
Claudia Solís-Román
Nicholas Commins
Livia Martinez

There is a large body of evidence showing that social-emotional learning (SEL) programs can boost short-term academic performance, social skills, and school attitudes.[1] These findings have encouraged policymakers and school districts to invest in SEL programs, especially because the COVID-19 pandemic significantly affected students from disadvantaged backgrounds. However, most research in this field still focuses on short-term results. One review found that 61 percent of SEL studies looked at effects for less than one year, and very few followed students through major life transitions, such as the move to middle or high school.[2] This gap is important because the benefits of early interventions often fade over time, a trend researchers call "fade-out."[3]

This long-term follow-up study of INSIGHTS into Children's Temperament (INSIGHTS) is unusual in that it has tracked the academic outcomes of students who participated in this two-year, temperament-based SEL program for kindergarten and first-grade students across the full span of their K-12 education. A companion brief, published in March 2026 found that, for the average student, early gains had faded by middle school, though students

mdrc
BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

from families with low incomes showed modest academic improvements.[4] This current brief provides the final chapter of the story, following students 13 years after they first joined the program in kindergarten through the year after their expected high school graduation. By using administrative data from New York City Public Schools and the National Student Clearinghouse, the study examines whether this well-implemented early SEL program changed these students' long-term educational paths.

If children learned how to handle frustration, how to keep trying when things get hard, and how to get along with others, these skills could help them persist and succeed in school long after the program is over.

What Is INSIGHTS into Children's Temperament (INSIGHTS)?

INSIGHTS is a distinctive early SEL intervention grounded in temperament theory.[5] Unlike universal SEL curricula that deliver identical content to all students, INSIGHTS uses a child-centered approach focused on "goodness of fit." This approach is based on the idea that children do their best when their specific traits match the demands of their school and home lives. When these traits do not align, children may struggle with schoolwork or behavior even if they are very capable. The program teaches adults how to recognize and respond to different temperaments. It also helps children understand their own traits using puppets that represent four different temperament profiles.

During kindergarten and first grade, INSIGHTS provided coordinated support in three different ways. Teachers attended 10 weekly sessions each year to learn how to recognize temperament traits and view challenging behaviors as expressions of a child's personality rather than simple defiance. A similar program for parents provided 10 sessions each year to help families use these same strategies at home. Finally, a classroom program used puppets in 10 weekly, 45-minute sessions to help students understand their differences and practice solving problems. The teacher and classroom components reached all students in the INSIGHTS schools, and parents that opted into the program received the parent portion.

Expecting a two-year program delivered in kindergarten and first grade to shape student outcomes a decade later is ambitious, and any such claim requires a clear theoretical rationale. Researchers suggest there are at least two probable ways this could happen.[6] First, the program might have improved early skills in reading and math in a way that builds on itself as a student grows. Early academic advantages can lead to stronger performance that builds across grades. Second, the self-regulation and social skills taught in INSIGHTS (which are not measured directly in this follow-up study) could be lasting abilities that students take with them into new environments. If children learned how to handle frustration, how to keep trying when things get hard, and how to get along with others, these skills could help them persist and succeed in school long after the program is over.

However, both of these ideas assume that future school environments do not wash away those early gains. As students progress through school, they encounter new teachers, complex social scenes,

and harder schoolwork. None of these later challenges are directly addressed by an intervention that only happens in the early grades. Lasting effects are possible but not guaranteed. In fact, most research on early intervention suggests that keeping early improvements alive usually requires continued support and reinforcement.[7]

The Original Trial and Past Findings

This follow-up study is based on a cluster-randomized trial that included 22 public elementary schools in New York City.[8] These schools were recruited in three groups between 2008 and 2010 and were randomly assigned to either the INSIGHTS program or a control group. In the control group, students took part in a supplementary reading program after school, while their parents and teachers attended workshops focused on literacy (referred to as the "attention-control" group hereafter). This specific design helped researchers ensure that any positive effects were due to the INSIGHTS lessons, not just the fact that students were receiving extra attention.

The current study followed 1,329 kindergarteners who were enrolled in the 22 study schools at the time of random assignment, with 641 in INSIGHTS schools and 688 in the attention-control schools. The sample was drawn from an urban, high-poverty area where over 70 percent of students were Black and about 75 percent were eligible for free or reduced-price lunch. The randomization successfully created two equal groups at the start, and the program was delivered as intended, with over 90 percent of the parent and teacher content completed.[9]

The original trial demonstrated substantial short-term benefits. Students in INSIGHTS schools whose parents opted into the program grew much faster in reading (effect size = 0.55) and mathematics (effect size = 0.31) during intervention years, and the program also boosted attention while reducing disruptive behaviors.[10] However, later studies showed that these academic effects diminished by fifth and sixth grade.[11]

The middle school follow-up study, which tracked the full randomized sample through eighth grade, confirmed this trend.[12] For the average student, there were no noticeable differences in English or math performance, grade retention, attendance, or special education status by the time they reached the end of middle school. For students from low-income families, the middle school results were more positive with modest improvements in English and math scores and reduced chances of a student being placed in special education. The findings for students from low-income families were exploratory and should be viewed with some caution because of the smaller sample size and the risk of finding patterns by chance.[13]

So far, the dominant pattern for the average student has been one of fade-out, where promising early gains slowly disappeared as the students got older.

What Is This Brief About?

This brief presents the final set of results from the long-term study of INSIGHTS. Using administrative records, it tracks all the students from their start in kindergarten through their expected high school graduation and their transition to college, 13 years later.

Because this study relies on administrative records, the outcomes it can examine are limited to academic and educational milestones. No direct measures of social-emotional skills were available for the high school years, meaning the study cannot speak to whether the social-emotional gains produced by INSIGHTS during the intervention period persisted as students aged.

The study focused on three prespecified primary outcomes to test whether INSIGHTS produced lasting effects:

- **NINTH-GRADE ON-TRACK STATUS:** This outcome measures whether students earned at least 10 course credits and passed at least one New York State Regents exam by the end of their expected ninth-grade year — a measure developed by the Research Alliance for New York City Schools that strongly predicts on-time graduation.[14]

- **RECEIPT OF A NEW YORK STATE HIGH SCHOOL DIPLOMA:** This outcome measures whether students received a New York State diploma by the end of their expected senior year.[15] The study also examined cumulative diploma receipt allowing one additional year, recognizing that some students, particularly those affected by COVID-19 disruptions, may have taken longer to complete their requirements.

- **IMMEDIATE COLLEGE ENROLLMENT:** This outcome measures whether students enrolled in a two-year or four-year college during the year immediately following their expected high school graduation, using data from the National Student Clearinghouse. The Clearinghouse tracks data from nearly 3,600 schools and covers over 97 percent of all college enrollments in the United States.[16] Given this near-universal coverage, students who do not appear in the database are coded as not enrolled rather than as missing.

The study also examined several exploratory outcomes to provide additional context for student engagement and behavior available from the administrative data: attendance rate, grade retention, and special education classification during high school, as well as course credit accumulation and Regents exam passage at each grade. Together, these different measures help paint a complete picture of what the students' high school years were like. In addition, the study explored the effects of INSIGHTS on students from low-income families and those whose parents consented to the study. All results are measured based on the grade level students were expected to be in, rather than their actual grade. This approach preserves the experimental design and avoids introducing bias, since the program itself might affect whether a student is held back. Robustness checks using other sample definitions are reported in the Technical Supplement.

The Study Sample

Table 1 outlines which grade level each group was expected to be in during each school year. The three cohorts of students entered kindergarten between 2008 and 2010 and were expected to reach their senior year between 2021 and 2023. By following students through the 2023-2024 school year, the study was able to get at least one full year of postsecondary data for all study participants.

Table 1. Summary of Follow-Up Grade and Year, by Study Cohort

Academic Year	Cohort 1	Cohort 2	Cohort 3	
2008-2009	Kindergarten	-	-	
2009-2010	1st grade	Kindergarten	-	
2010-2011	2nd grade	1st grade	Kindergarten	**Program implementation period**
2011-2012	3rd grade	2nd grade	1st grade	
2012-2013	4th grade	3rd grade	2nd grade	
2013-2014	5th grade	4th grade	3rd grade	
2014-2015	6th grade	5th grade	4th grade	
2015-2016	7th grade	6th grade	5th grade	
2016-2017	8th grade	7th grade	6th grade	
2017-2018	**9th grade**	8th grade	7th grade	
2018-2019	**10th grade**	**9th grade**	8th grade	
2019-2020	**11th grade**	**10th grade**	**9th grade**	**COVID-19 period**
2020-2021	**12th grade**	**11th grade**	**10th grade**	
2021-2022	**Post 1**	**12th grade**	**11th grade**	
2022-2023	Post 2	**Post 1**	**12th grade**	
2023-2024	Post 3	Post 2	**Post 1**	

NOTES: Bolded grades are those covered in this brief. This table describes the expected year and grade for students following an educational trajectory with no grade retention and no accelerated promotion (that is, skipping a grade).

All three cohorts were in high school when the COVID-19 pandemic began, though they were at different expected grade levels. Cohort 1 students were expected to be in eleventh grade when schools closed in spring 2020; Cohort 2 students were in tenth grade; and Cohort 3 students were in ninth grade. The pandemic interrupted learning, attendance, and testing processes. For the oldest students, it also disrupted the college application process. Although the study design ensures that both groups were affected equally, the pandemic is an important factor to consider when looking at graduation and college enrollment rates and understanding the environment in which postsecondary decisions were made.

Like most long-term studies in large urban school districts, the number of participants dropped over time as families moved or students left the New York City school system. After 13 years, about 70 percent of the original group was still enrolled in the city's schools or had recent records. The proportion of students who left the school system was about the same for both groups, and among students who remained, the two groups continued to look alike on baseline characteristics. These analyses are reported in full in the Technical Supplement (see Tables S.1 and S.2).

High School Follow-Up Study Findings

The research team assessed the impacts of INSIGHTS on high school and postsecondary outcomes using the full analytic sample of students enrolled in study schools at random assignment. The findings are organized around three questions: Did the early benefits of the INSIGHTS program persist through the expected high school years and beyond? Can the unexpected college enrollment findings be explained? Did effects differ for the key subgroups that were identified in the middle school follow-up study?

Finding 1: Early Benefits Did Not Persist Through High School for the Average Student

Across the measures of high school progression and completion examined in this study, INSIGHTS produced no detectable impacts for the average student. Students in the INSIGHTS and attention-control groups progressed through school at nearly identical rates. They earned similar numbers of credits, passed roughly the same number of Regents exams, attended school at the same frequency, and graduated at similar rates. This result confirms the fade-out pattern seen in the middle school follow-up study.

Table 2 shows that students in both groups were equally likely to be on track at the end of ninth grade: 51.7 percent of INSIGHTS group students compared with 51.9 percent of attention-control group students. The proportion of students who received a New York State diploma by the end of their expected senior year was also virtually identical—37.4 percent for INSIGHTS group students and 37.1 percent for attention-control group students. The findings remained the same when allowing one additional year to complete diploma requirements. None of the differences between the two groups are statistically significant.

Other measures of high school progress further supported these results. Through the eleventh grade, students in both groups earned almost the same number of credits and passed a similar number of Regents exams at each measurement point. Attendance rates were also comparable at roughly 81 or 82 percent, and students in both groups were held back at similar rates of about 13 percent. While about 5 percentage points fewer INSIGHTS students were placed in special education compared with the attention-control group (29.2 percent versus 34.1 percent), this difference was not statistically significant.

Table 2. Estimated Impacts of INSIGHTS on High School and Postsecondary Outcomes for the Full Sample, Based on Expected Grade Level

Outcome	INSIGHTS Group	Attention-Control Group	Estimated Impact	Standard Error	P-Value
Academic progression and behavior during high school					
Ninth-grade on-track status (%)	51.7	51.9	-0.2	4.0	0.964
Cumulative number of credits earned					
Expected grade 9	9.00	9.18	-0.19	0.54	0.733
Expected grade 10	20.31	20.38	-0.07	0.88	0.938
Expected grade 11	32.01	31.57	0.43	1.31	0.744
Cumulative number of Regents tests passed					
Expected grade 9	2.06	2.32	-0.25	0.22	0.253
Expected grade 10	3.96	4.06	-0.09	0.31	0.772
Expected grade 11	5.57	5.65	-0.08	0.35	0.821
Average high school attendance rate (%)	81.9	80.7	1.2	1.6	0.442
Ever retained in grades 9 through 12 (%)	13.3	12.6	0.6	2.6	0.810
Ever identified for special education in grades 9 through 12 (%)	29.2	34.1	-4.8	3.3	0.158
Receipt of a New York State high school diploma					
On time[a] (%)	37.4	37.1	0.4	3.6	0.918
With a one-year delay (%)	46.3	48.7	-2.4	3.2	0.477
Immediate college enrollment[b] at					
Any postsecondary institution (%)	26.2	32.2	-6.0	3.6	0.117
Four-year postsecondary institution (%)	21.2	27.3	-6.1*	3.3	0.084
Two-year postsecondary institution (%)	5.6	4.9	0.7	1.5	0.663

SOURCE: District records collected between 2007 and 2024 and provided by the Research Alliance for New York City Schools.

NOTES: Outcomes are measured according to the follow-up year (the grade level a student is expected to be in).

[a]On time = received a New York State high school diploma by the end of the student's expected twelfth-grade year (13 years after starting kindergarten in the study).

[b]Immediate college enrollment = postsecondary enrollment in the school year immediately following the student's expected high school graduation year.

Total sample size varies by outcome, ranging from 831 to 1,329.

Values for the INSIGHTS group are the simple means for each outcome. Impact values are estimated from a hierarchical regression that controls for cohorts of random assignment and school and student level baseline covariates. Values for the attention-control group equal the values for the INSIGHTS group minus the estimated impacts.

Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed t-test was applied to the estimated difference. Statistical significance levels are indicated as: * = $p < 0.10$; ** = $p < 0.05$; *** = $p < 0.01$.

Overall, these results show that the early benefits of the program did not lead to noticeable differences by the time students reached the expected end of high school.[17] This finding matches the fade-out pattern seen in other early childhood programs and it highlights how difficult it can be to maintain early benefits through the teenage years without continued support.[18]

The only outcome that did not follow this pattern was college enrollment. Students in the INSIGHTS group were about 6 percentage points less likely to enroll in college the year after their expected high school graduation, at 26.2 percent compared with 32.2 percent for the attention-control group. This difference fell just short of conventional statistical significance ($p = 0.117$) but was concentrated in enrollment at four-year institutions, where the gap was 6.1 percentage points ($p = 0.084$). Two-year college enrollment showed no difference between the groups. This finding is surprising because it contradicts the program's goals and its early success. Since no other steps in the expected high school years showed this negative trend, the research team conducted multiple analyses to try and explain the result. Those investigations are detailed below.

Finding 2: An Exploratory Investigation of the College Enrollment Pattern Could Not Identify a Clear and Definitive Explanation

As mentioned earlier, the results for college enrollment are hard to explain given the rest of the evidence. The INSIGHTS program did not show any negative effects on intermediate steps like graduation, course credits, test scores, or attendance. In all of these areas, the two groups were essentially the same. This raises a difficult question: How could a program from early elementary school reduce college enrollment without negatively affecting any step along the pathway? It is also unclear why a program that improved reading, math, and self-regulation in the early grades would lead to a negative result 10 years later. To make sense of this, the research team investigated three potential factors.

Could the Timing or the Impact of COVID-19 Be the Cause?

College enrollment was measured in the year right after a student's expected on-time high school graduation. For all three cohorts of students in the study, this window occurred after the onset of the COVID-19 pandemic, a period of major disruption that included remote learning, economic stress, and changes to how colleges accepted students. However, for COVID-19 to explain the negative college enrollment pattern, it would have had to affect INSIGHTS group students differently than it affected attention-control group students. Because both groups attended New York City public schools in similar neighborhoods and faced the same pandemic conditions, there is no clear reason to expect a differential impact. If anything, the self-regulation skills emphasized by INSIGHTS might have helped students navigate disruption more effectively, which would work against the pattern observed here.

As an additional check, the research team compared average outcomes across the three study cohorts of students who experienced the pandemic at different points in their high school careers (Technical Supplement Table S.5). If pandemic disruptions to students' later high school years —

when they would have typically been taking standardized tests and applying to college — were driving the result, the negative pattern should be strongest for the first cohort, whose expected senior year coincided with the height of the pandemic (2020-2021 school year). Instead, that cohort showed almost no difference in college enrollment between the two groups. The negative trend appeared in the second and third cohorts of students, whose senior years were farther away from the pandemic's initial disruptions. These cohort comparisons involve small numbers of schools and should be interpreted with caution, but they provide additional evidence that COVID-19 is unlikely to be the primary explanation.

Longer enrollment windows, which could reveal whether INSIGHTS students enrolled later rather than not at all, could not be reliably assessed because not enough follow-up data are available for all three groups beyond the first year.

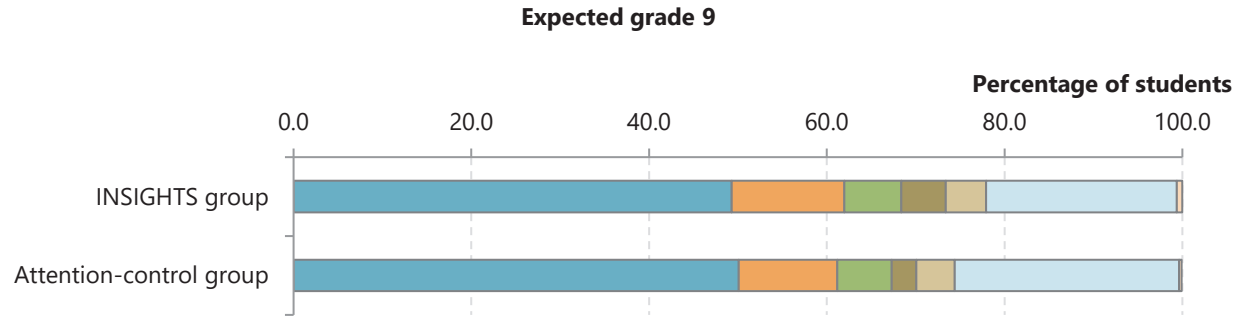
Could the Results be Due to Who Remained in the Study or How Outcomes Were Measured?

Over the course of 13 years, some students naturally left the New York City public school system. If the types of students who remained in the INSIGHTS group were different from those in the attention-control group, it could lead to incorrect conclusions. To test this, the team looked at which grades students were in and the types of schools they attended throughout their expected high school years.

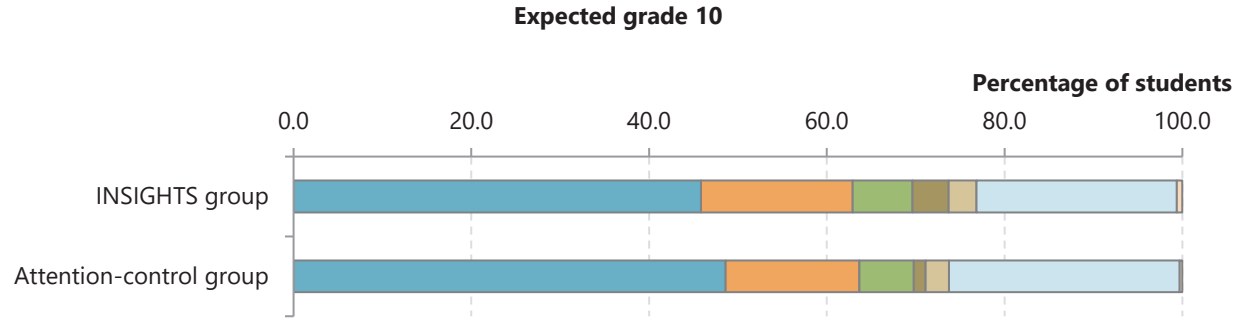
As Figure 1 shows, the INSIGHTS and attention-control groups looked similar throughout high school in terms of the type of schools they attended (regular public school versus charter school) and their grade progression (at or above expected grade level versus behind grade level). At each grade, comparable proportions of students were on track, behind grade level, enrolled in charter schools, or discharged from the system. The only statistically significant and consistent difference was that INSIGHTS students in charter schools were slightly more likely to be behind grade level, but this group is small with only some 50 students.

The research team further tested impacts under several alternative sample definitions to see whether sample composition could explain the college enrollment finding (Technical Supplement Table S.6). First, the team examined outcomes based on students' actual progression through high school rather than their expected grade level. This approach measures outcomes based on the years from when a student first enrolled in ninth grade in a New York City public high school. For example, it defines "on-time graduation" as earning a diploma within four years of starting ninth grade, and "immediate college enrollment" as enrolling in a college in the fifth year following ninth-grade entry. Using this definition, the negative effect on college enrollment was even larger. INSIGHTS group students were 10.8 percentage points less likely to go to college than attention-control group students ($p = 0.025$). Again, this was mostly because fewer students enrolled in four-year colleges (9.5 percentage points, $p = 0.057$). Meanwhile, there was still no significant difference in graduation rates between the two groups.

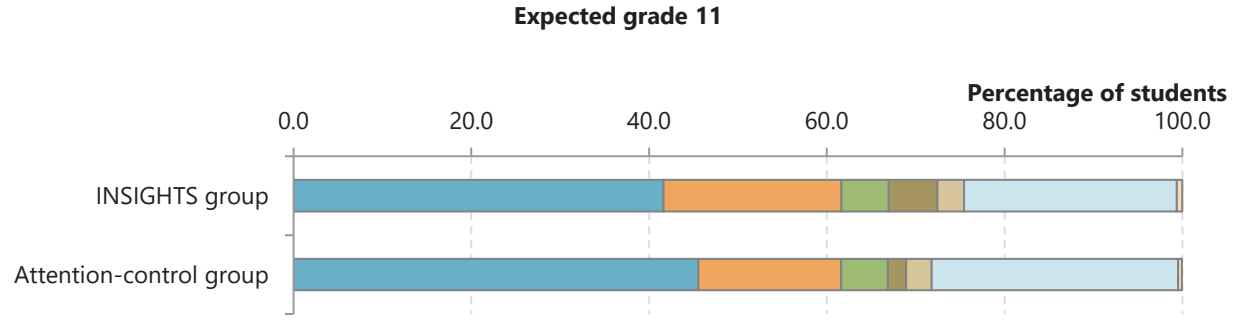
Figure 1. Where Were Students Enrolled at Each Expected Grade?



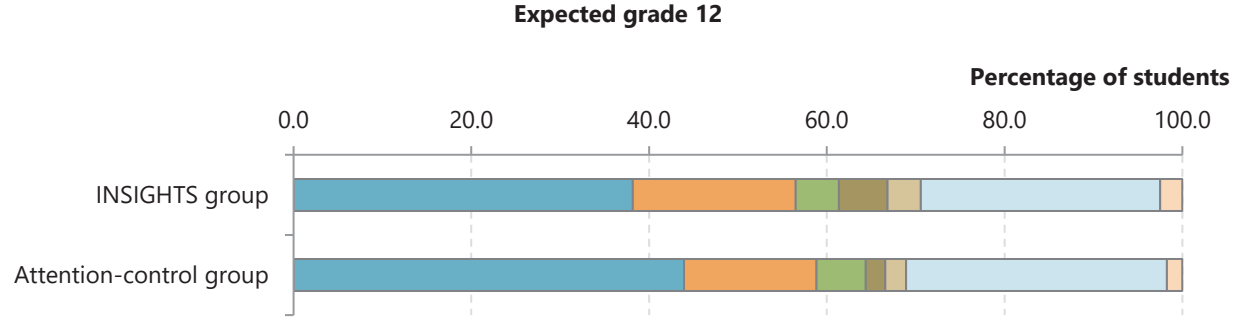
Significant differences: charter, below grade (*); discharged (*).



Significant differences: charter, below grade (***).



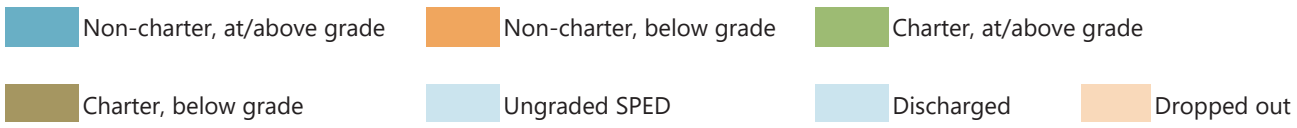
Significant differences: charter, below grade (***); discharged (*).



Significant differences: charter, below grade (**).

(continued)

Figure 1 (continued)



SOURCE: District records collected between 2007 and 2024 and provided by the Research Alliance for New York City Schools.

NOTES: The INSIGHTS group includes 641 students. The attention-control group includes 688 students. Stacked bars show the percentage of the full original sample (as randomized at kindergarten entry) in each category at each expected high school grade. Categories are based on students' enrollment records, actual grade level relative to expected grade level, charter school status, and special education classification.

All grades reflect the grade levels students were expected to be in for a given program year. "Non-charter at/above grade" and "Charter at/above grade" include students enrolled at or above their expected grade level. "Below grade" categories include students enrolled below their expected grade level. "Ungraded SPED" includes students receiving special education services who are not classified by traditional K-12 grade levels. "Discharged" includes students who left the New York City public school system. "Dropped out" includes students with a confirmed dropout discharge code. Percentages may not sum to exactly 100 because of rounding.

A two-tailed t-test was applied to each estimated difference. Statistical significance levels are indicated as: * = $p < 0.10$; ** = $p < 0.05$; *** = $p < 0.01$.

Restricting the analysis to students who were actively enrolled in a New York City public school or to students who were actively enrolled in the twelfth grade during their expected senior year did not change the negative college enrollment pattern either. In many cases, these alternative ways of looking at the data made the negative pattern even clearer. These results suggest that the college enrollment findings were not caused by the timing of the data, the types of students who stayed in the study, or whether students were held back.

Could the Schools the Students Attended in the Follow-Up Period Be Responsible?

The most theoretically grounded explanation connects to INSIGHTS' central mechanism — goodness of fit between children and their environments. If INSIGHTS and attention-control group students ended up in systematically different schools during the years after the intervention, those environmental differences could affect long-term outcomes. Schools can differ in many ways, including their safety climate, academic expectations, college-going culture, quality of teacher-student relationships, and emphasis on SEL. If INSIGHTS students attended schools that were, on average, less supportive in ways that matter for college enrollment, it could explain the results even if INSIGHTS itself had no direct effect.

The team investigated this hypothesis using two complementary approaches. First, the team linked student records to annual school-level learning environment measures from the New York City School Surveys (formerly the Learning Environment Surveys) and examined whether the two groups attended schools with different environments across eight domains, including safety and respect,

academic expectations, teacher-student relationships, SEL practices, and college and career readiness culture, among others.[19]

As shown in Table 3, students from both groups attended schools with broadly similar environments. Where differences did appear, they were small and scattered across domains, with no clear or consistent pattern favoring either group. Of 50 available domain-level contrasts examined across eight domains and seven grades, only four (8 percent) were statistically significant at the 10 percent level — a rate consistent with what would be expected by chance alone. The significant differences were scattered across different domains and years with no clear or consistent pattern. More detailed analyses that broke down each domain into its individual components produced similar results: 13 of 129 estimated contrasts (about 10 percent) reached significance, again with no coherent pattern. Most importantly, there were no consistent patterns across years in the areas most closely related to the program's theory of change, such as teacher-student relationships, SEL practices, and safety and respect (Technical Supplement Table S.9). The schools' focus on college and career readiness also showed no consistent differences between the two groups.

Table 3. Estimated Contrasts on School Environment Survey Constructs Between INSIGHTS Group and Attention-Control Group Students, by Domain and Expected Grade

New York City School Survey Composite	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Teacher-student relationships	-0.05	-0.05	0.03	0.06	0.01	-0.03	-0.05
Social-emotional learning practices	-0.06	-0.08	0.05	0.15**	0.01	-0.05	0.02
Safety and respect	-0.13	-0.18*	-0.10	-0.08	0.02	-0.11**	-0.07
College and career readiness culture	—	—	—	0.00	-0.08	-0.08	-0.07
Academic expectations	0.08	-0.05	0.03	0.03	-0.03	-0.04	-0.11**
Student engagement	-0.07	-0.07	0.00	0.07	—	—	—
School leadership	0.00	0.08	0.02	0.07	0.02	-0.01	0.05
Culturally responsive practices	-0.10	-0.14	-0.05	0.05	-0.04	0.02	-0.06

SOURCE: District records collected between 2007 and 2024 and provided by the Research Alliance for New York City Schools; NYC Public Schools, NYC School Surveys, 2009-2024.

NOTES: Each cell shows the estimated difference (in standard deviation units) between the school environment domains of schools attended by INSIGHTS students and schools attended by attention-control group students at a given expected grade. (The domains are made up of all the subconstructs within those categories.) Positive values indicate that INSIGHTS students attended schools with higher scores on that domain; negative values indicate lower scores. Estimates are from a hierarchical regression that controls for cohorts of random assignment and school- and student-level baseline covariates.

Domain composites are the mean of constituent subconstruct z-scores, subject to a 50 percent minimum item threshold. Internal consistency was high: Cronbach's alpha for domain-level composites and subconstructs within domains are reported in Technical Supplement Table S.8.

A dash (—) indicates that the measure is not applicable for schools for that grade (for example, graduation rates are not reported for middle schools), that data were not available, or that valid data were available for fewer than 50 percent of the sample for the grade.

A two-tailed t-test was applied to each estimated difference. Statistical significance levels are indicated as: * = $p < 0.10$; ** = $p < 0.05$; *** = $p < 0.01$.

Second, the research team examined objective school characteristics drawn from the New York City Public Schools' School Quality Reports (SQR).[20] These reports provide annual data on New York City public schools, including who attends the school, how students perform academically, and where graduates go after high school. From these reports, the team selected measures in three categories chosen for their relevance to the study's long-term outcomes: a school's college-going environment (the share of graduates who enrolled in postsecondary education, the school's college and career preparedness index, and the school's career and technical education designation); its academic performance (graduation rates, student persistence rates, and the share of students who were overage and under-credited); and the demographic composition of its student body (economic need, enrollment size, and racial/ethnic makeup).[21] Unlike the survey-based climate measures, which capture how students and teachers experience their schools, these characteristics capture the measurable profile of the institution itself.

As shown in Table 4, the comparison of these school characteristics produced a more mixed picture than the survey-based analysis. Most student body composition measures showed no significant differences: INSIGHTS and attention-control group students attended high schools of similar size and with similar proportions of English language learners and students with individualized education programs. However, INSIGHTS students attended schools with a modestly higher economic need index and a somewhat lower share of White students, differences that were consistent in direction across grades though small in magnitude.

The more notable pattern appeared in measures related to school performance and college enrollment. During expected grades 9 through 11, INSIGHTS students attended high schools where the share of graduates enrolling in postsecondary education within 6 and 18 months was roughly 3 percentage points to 5 percentage points lower than the share at schools attended by attention-control group students. The high schools attended by INSIGHTS students also had four-year graduation rates that were about 2 percentage points lower, and a somewhat larger share of students who were overage and under-credited. These differences were statistically significant at multiple grade levels and consistent in direction, though they narrowed somewhat by the expected senior year.

Notably, this pattern is consistent with small differences observed at the start of the study. Although randomization successfully created equivalent groups of students, the elementary schools in the INSIGHTS group had a modestly higher concentration of students from low-income families and a somewhat larger share of Black students compared with attention-control group schools.[22] These baseline differences were not statistically significant but are directionally consistent with the high school-level differences documented here, suggesting that the pattern may reflect pre-existing geographic or neighborhood sorting rather than being a consequence of the intervention itself. Future research could investigate whether school feeder patterns or neighborhood-level factors help explain how students in both groups end up in somewhat different school environments later.

Table 4. Estimated Differences in Selected School Characteristics Between Schools Attended by INSIGHTS Group and Attention-Control Group Students, by Expected High School Grade

School Characteristics	Grade 9	Grade 10	Grade 11	Grade 12
College-going environment				
College and career preparedness index	-0.04	0.00	-0.02	0.02
Postsecondary enrollment rate, 6 months (%)	-5.1***	-3.6*	-3.5**	-1.6
Postsecondary enrollment rate, 18 months (%)	-4.1**	-3.4**	-3.3*	-3.1**
Attended CTE schools (%)	2.2	3.7	3.7	3.7
School performance				
Over 90% attendance rate (%)	-5.4*	—	—	—
Overage and under-credited students (%)	1.7***	1.4***	1.9**	-0.3
Grade 9 students earning more than 10 credits (%)	-0.9	-1.8	-1.7	-1.6
Average Regents score, Algebra I	—	—	—	-0.65
Average Regents score, English	—	—	—	-1.03
Four-year high school persistence rate (%)	-1.7**	-1.3**	-1.3**	-0.7
Six-year high school persistence rate (%)	-1.4	-1.2*	-1.3*	-1.7***
Four-year graduation rate (%)	-2.3**	-1.8**	-2.3***	-1.6*
Six-year graduation rate (%)	-2.1**	-1.6**	-1.7**	-2.0***
Student body composition				
Enrollment	-12.20	3.93	25.67	17.24
Economic need index	0.04**	0.04*	0.04	0.03*
English language learners (%)	0.8	0.6	0.6	0.9
Students with Individualized Education Programs (%)	—	—	0.3	-0.1
Black, non-Hispanic (%)	5.1	4.3	3.7	3.1
Hispanic (%)	-1.6	-1.0	-0.6	0.0
White, non-Hispanic (%)	-3.6*	-3.2*	-3.2*	-3.0*

SOURCES: District records collected between 2007 and 2024 and provided by the Research Alliance for New York City Schools; New York City Public Schools, School Quality Reports, 2014-2015 through 2023-2024.

NOTES: CTE = Career and technical education. School characteristics are drawn from the New York City Public Schools School Quality Reports and reflect the reported values for the school a student attended in a given year. For students enrolled in different schools at the October and June snapshots, the two values were averaged.

Estimates are from a hierarchical regression that controls for cohorts of random assignment and school- and student-level baseline covariates. Student composition measures (which are available for both middle and high schools) are shown for expected grades 6 through 12. School performance and college-going measures (which are primarily available for high schools) are shown for expected grades 9 through 12 where data permit.

A dash (—) indicates that the measure is not applicable for schools for that grade (for example, graduation rates are not reported for middle schools); that data were not available; or that valid data were available for fewer than 50 percent of the sample for the grade.

A two-tailed t-test was applied to each estimated difference. Statistical significance levels are indicated as: * = $p < 0.10$; ** = $p < 0.05$; *** = $p < 0.01$.

These patterns are suggestive, but several important caveats apply. The observed differences are small in absolute terms. A 2 percentage point to 3 percentage point difference in a school's graduation rate or a 3 percentage point to 5 percentage point difference in its college-going rate is meaningful at the school level but difficult to connect directly to the 6 percentage point gap in the likelihood of individual-level college enrollment between the two groups. The data underlying some measures had substantial missing values, particularly high school-specific variables in the earlier grades, and the analysis was limited to measures with valid data for at least half of the sample. The team also examined a large number of school characteristics across multiple grades. Of 69 contrasts shown in Table 4, 32 were statistically significant at the 10 percent level, a rate that exceeds what would be expected by chance but should still be interpreted with caution given the volume of comparisons. Full results across all available measures and expected middle school and high school grades are reported in Technical Supplement Table S.10.

Finally, even if INSIGHTS students did attend somewhat different high schools, it is not clear that the intervention itself influenced later school choices. The elementary schools in the INSIGHTS group already had modestly different demographic profiles at the start of the study, and school sorting after elementary school is shaped by many factors, including geographic proximity, feeder patterns between school levels, family preferences, and residential moves. The high school differences observed here may partly reflect where students started rather than being a consequence of participating in INSIGHTS, though the possibility that the program shaped students' or families' orientations toward certain types of schools cannot be ruled out. Regardless of the source of these differences, attending high schools with somewhat lower college-going rates could have contributed to the enrollment gap, but the differences are too small and uncertain to serve as a full and definitive explanation.

What Remains Unexplained

These investigations narrowed the field of possible explanations without arriving at a definitive answer. For COVID-19 to explain the pattern, it would have had to affect INSIGHTS students differently than attention-control group students, which is unlikely since both groups attended New York City public schools in similar neighborhoods and faced the same pandemic conditions. Survey-based measures of school climate showed no systematic differences between the schools the two groups attended. Objective school characteristics told a more nuanced story: INSIGHTS students attended high schools with modestly lower college-going and graduation rates and slightly higher concentrations of economic disadvantage. These differences are directionally consistent with the college enrollment finding and consistent with insignificant demographic differences between the two groups' elementary schools at the start of the study. However, the differences are small, accompanied by substantial data limitations, and do not establish a causal link. It is not clear whether they reflect pre-existing geographic or neighborhood sorting, or some influence of the program on families' later school choices, or both.

The negative college enrollment pattern could not be clearly attributed to COVID-19 disruptions, differences in who remained in the study, grade retention, or systematic differences in either the climate or characteristics of the schools students attended. The finding remained consistent, and

in some cases grew stronger, across multiple analytical model specifications, estimation methods, and sample definitions (see Technical Supplement Tables S.4 through S.6).

At the same time, there are several reasons to be careful when interpreting the college enrollment finding. The 6 percentage point difference in the primary specification did not reach conventional statistical significance. The study examined a large number of outcomes across grades, subgroups, and specifications, and some unexpected results are inevitable when many comparisons are made. Also, the measurement window for college enrollment captures only the first year after expected graduation. Many students may still choose to go to college later on. Most importantly, the finding contradicts both the program's earlier positive effects and its theory of change: There is no logical mechanism through which a kindergarten intervention focused on temperament awareness and classroom fit would reduce college enrollment a decade later without affecting any outcome in between.

For these reasons, this should be seen as a finding that deserves further investigation, not as a cause for alarm. The analysis of school characteristics identified a thread worth pursuing—the possibility that school sorting after elementary school, whether driven by pre-existing neighborhood patterns or other factors, plays a role—but it remains only a hypothesis. Future research with longer follow-up windows, larger samples, more detailed data on school selection processes and career and technical education pathways, and direct measures of college decision-making could help determine whether this pattern replicates and, if so, what drives it.

Finding 3: Subgroup Analyses Showed That Early Benefits Did Not Last Through High School

The middle-school follow-up study identified two student groups of interest. These included students from low-income backgrounds (defined by their baseline free or reduced-price lunch eligibility status), who showed modest academic gains and lower special education rates through eighth grade, and students whose parents took part in the parent sessions of the program. This follow-up study looked to see whether these groups fared better than the overall sample in high school. They did not.

Students from low-income backgrounds did not show significant improvements in ninth-grade on-track status, graduation, or college enrollment (see Technical Supplement Table S.11). However, one exploratory result did show a 7.9 percentage point reduction in special education placement during high school for this group (p-value = 0.048), which continues a trend first noticed in middle school. This result was sensitive to how special education status was measured, so it should be viewed with some caution.[23]

Students whose parents consented to the study also showed no clear long-term benefits in high school (see Technical Supplement Table S.12). The estimates for on-time diploma receipt favored the INSIGHTS group by 9 percentage points, but this difference was not statistically significant (p-value = 0.122). This result also did not persist when allowing an additional year for graduation

and was not supported by corresponding differences in ninth-grade on-track status, credit accumulation, or Regents exam passage during earlier high school years. College enrollment for this subgroup followed the same negative direction as the full sample, though the estimates were smaller and not statistically significant. Since more graduates should mean more students positioned to enroll in a postsecondary institution, the disconnect between the suggestive diploma advantage and the negative college enrollment direction suggests the graduation estimate may reflect sampling variability in a smaller subgroup (N = 394) rather than a genuine program effect.

These findings reinforce the study's main conclusion. The early benefits of the INSIGHTS program did not last through the end of high school, even for groups that showed more promise earlier on. The one sustained benefit observed in middle school for this subgroup — reduced special education classification — may have continued into high school, but the evidence is not robust enough to draw firm conclusions.

What These Findings Mean for Policy and Practice

This study provides some of the longest follow-up evidence available on an early-grade SEL intervention, tracking students from kindergarten through expected high school graduation and into the first year of postsecondary education. Because INSIGHTS is a temperament-based program delivered in a specific urban context and compared with an attention-control reading program, the findings should be interpreted as evidence about this particular intervention and similar early elementary approaches, not as a verdict on SEL programs broadly. With that caveat, the results offer several insights for policymakers, school district leaders, and practitioners.

- **TWO YEARS OF SEL PROGRAMMING IN THE EARLY GRADES WERE NOT ENOUGH TO CHANGE LONG-TERM EDUCATIONAL TRAJECTORIES ON THEIR OWN.** The fade-out pattern seen in this study is remarkably consistent across different groups and outcomes. While the program was running, it produced real gains in reading, math, and behavioral outcomes. But those gains slowly disappeared during elementary school. By middle school, those gains were mostly gone, and they left no noticeable mark on high school diploma receipt or college enrollment for the average student. This finding does not mean the early gains were without value as they likely improved students' daily lives and classroom experiences during an important period in their development. However, the finding does suggest that a short-term intervention in the early grades, even one that is well designed and implemented, may not lead to longer-term changes in academic outcomes if it is not followed by continued support.

  This finding aligns with what most SEL frameworks recommend. Organizations such as the Collaborative for Academic, Social, and Emotional Learning emphasize that social-emotional development is an ongoing process that benefits from reinforcement across all grade levels.[24] The theory of change for INSIGHTS articulated plausible pathways through which early gains in self-regulation and academic skills could cascade into longer-term benefits, but the program itself was designed to operate in kindergarten and first grade. These results suggest that the

early skills the program built were not sufficient on their own to carry students through the very different demands of middle and high school without additional reinforcement. The findings should not be read as evidence that early SEL is ineffective. Rather, these results underscore that early investments may need to be paired with developmentally aligned programming in later grades, particularly during key transitions, when research consistently documents declines in motivation and engagement, to sustain their benefits over time.[25]

- **SCHOOLS INTERESTED IN SUSTAINING EARLY GAINS SHOULD CONSIDER ALIGNING THEIR SEL APPROACHES ACROSS GRADES.** The core idea of INSIGHTS is that children thrive when their environment is responsive to their individual characteristics. This idea is just as important for older students as it is for younger ones. However, the specific strategies taught in INSIGHTS were designed for very young children. As students entered middle and high school, they encountered settings with different expectations, structures, and social demands. Without a clear effort to reinforce temperament-responsive practices in later grades, the skills students learned early on may have faded, though this study lacks direct measures of social-emotional skills in the later years to confirm whether those early gains persisted, diminished, or evolved. Districts that invest in early SEL programming might consider how those approaches connect to the systems students encounter in later years, whether through transition support, teacher professional development that spans grades, or explicit alignment of SEL goals across elementary, middle, and high school.

- **THE UNEXPECTED COLLEGE ENROLLMENT FINDING WARRANTS FURTHER INVESTIGATION, NOT CHANGES IN PRACTICE.** The negative college enrollment pattern is puzzling and, despite extensive investigation, has not been definitively explained. Exploratory analyses did reveal that INSIGHTS students attended high schools with modestly lower college-going and graduation rates. This pattern is directionally consistent with the enrollment gap and with demographic differences between the two groups' elementary schools at the start of the study, but these differences are small and do not establish a causal link. The finding should not, at this stage, alter decisions about whether to implement INSIGHTS or similar early SEL programs, particularly since the difference in enrollment in the year immediately following expected graduation did not reach conventional statistical significance in the primary specification. Future research with longer follow-up windows, larger samples, more detailed data on how families choose schools and on career and technical education pathways after elementary school, and direct measures of college decision-making could help determine whether this pattern replicates and what drives it.

- **LONG-TERM FOLLOW-UP STUDIES ARE ESSENTIAL FOR UNDERSTANDING EARLY INTERVENTIONS.** Short-term evaluations of INSIGHTS would have concluded that the program was highly effective. Only by tracking students beyond the intervention window did the fade-out pattern become visible, and notably, signs of that pattern emerged as early as the end of elementary school. Even follow-up studies of modest duration can reveal important information about whether early gains persist. Extending the follow-up period through the full K-12 arc, as this study did, provided additional value by capturing high school completion and college enrollment rates. These milestones matter greatly for students' long-term well-being but require patience and sustained investment

in longitudinal data infrastructure to measure. The field needs more evidence of this kind, not less, to make informed decisions about where and how to invest in children's development.

Looking Ahead

Thirteen years after the program began, students who participated in INSIGHTS were no more or less likely to receive a high school diploma than their peers. The early academic advantages they once had have faded. The unexpected college enrollment pattern raises questions that this study could not resolve and that merit continued attention. What is clear is that the path from a promising kindergarten program to lasting changes in academic trajectories is long and complex, and extends beyond the reach of any single intervention. Supporting students across that full trajectory, not just at its beginning, remains the central challenge for educators and policymakers committed to closing opportunity gaps.

Notes and References

1. Joseph A. Durlak, Roger P. Weissberg, Allison B. Dymnicki, Rebecca D. Taylor, and Kriston B. Schellinger, "The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions," *Child Development* 82, 1 (2011): 405–432.
2. Rebecca D. Taylor, Eva Oberle, Joseph A. Durlak, and Roger P. Weissberg, "Promoting Positive Youth Development Through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-Up Effects," *Child Development* 88, 4 (2017): 1,156–1,171.
3. Drew H. Bailey, Greg J. Duncan, Candice L. Odgers, and Winnie Yu, "Persistence and Fadeout in the Impacts of Child and Adolescent Interventions," *Journal of Research on Educational Effectiveness* 10, 1 (2017): 7–39.
4. Pei Zhu, Livia Martinez, Nicholas Commins, Claudia Solís-Román, and Peyton Nash, "Do Early Social-Emotional Learning Programs Have Lasting Effects Through Middle School? Middle School Follow-Up Findings from a Randomized Trial of INSIGHTS" (MDRC, 2026).
5. Sandee G. McClowry, *Temperament-Based Elementary Classroom Management* (Rowman & Littlefield, 2014).
6. For a framework describing three mechanisms for long-term effects of early interventions (skill accumulation, key transition points, and environmental changes), see Bailey, Duncan, Odgers, and Yu (2017). See also Drew H. Bailey, Greg J. Duncan, Flávio Cunha, Barbara R. Foorman, and David S. Yeager, "Persistence and Fade-Out of Educational-Intervention Effects: Mechanisms and Potential Solutions," *Psychological Science in the Public Interest* 20, 1 (2020): 8–37.
7. Bailey, Duncan, Odgers, and Yu (2017); Bailey, Duncan, Cunha, Foorman, and Yeager (2020).
8. In a cluster-randomized controlled trial, clusters (schools in this case) are randomly assigned either to a program group that is eligible to participate in the intervention or to a control group that is not eligible to participate in the intervention. By comparing the outcomes of the two groups, which are not systematically different in any way, a study can estimate the impact of the intervention without bias.
9. See Table 1 in Zhu, Martinez, Commins, Solís-Román, and Nash (2026) for baseline equivalence between the INSIGHTS and attention-control groups at both school and student levels. For an assessment of implementation fidelity, see Erin E. O'Connor, Elise Cappella, Meghan P. McCormick, and Sandee G. McClowry, "An Examination of the Efficacy of INSIGHTS in Enhancing the Academic and Behavioral Development of Children in Early Grades," *Journal of Educational Psychology* 106, 4 (2014): 1,156–1,169.
10. Estimated effect sizes are from O'Connor, Cappella, McCormick, and McClowry (2014). Effect sizes summarize the average difference between program and comparison groups in standard deviation units.
11. Meghan McCormick, Robin Neuhaus, Erin E. O'Connor, Hope I. White, E. Parham Horn, Samantha Harding, Elise Cappella, and Sandee McClowry, "Long-Term Effects of Social-Emotional Learning on Academic Skills: Evidence from a Randomized Trial of INSIGHTS," *Journal of Research on Educational Effectiveness* 14, 1 (2021): 1–27.
12. Zhu, Martinez, Commins, Solís-Román, and Nash (2026).
13. Subgroup analyses of this kind are considered exploratory because they were not the primary focus of the study and involve a smaller subset of the sample. When multiple comparisons are conducted across subgroups, there is a greater chance that some findings appear statistically significant purely by chance, in other words, they reflect random variation rather than true effects of the program. These findings should therefore be interpreted with caution and would benefit from replication in future studies.

14. James Kemple, Michael Segeritz, and Nickisha Stephenson, "Building On-Track Indicators for High School Graduation and College Readiness: Evidence from New York City" (Research Alliance for New York City Schools, 2013). New York State Regents exams are standardized tests administered by the New York State Education Department covering core high school subjects including English, mathematics, science, and social studies. Students must pass a minimum number of Regents exams to earn a New York State high school diploma.
15. The graduation measure draws from the Research Alliance for New York City Schools graduation cohort file, which includes all New York City public school students who received a New York State diploma. To preserve the original randomization sample, students who were discharged from the school system, who could not be located in the data, or who did not have a graduation record are coded as not receiving a diploma rather than as missing. This approach maintains the integrity of the experimental comparison but may lower overall diploma receipt rates relative to what would be observed in a typical graduating class. Sensitivity analyses using alternative coding methods, such as treating discharged students as missing rather than as not graduating, yielded similar results, confirming that the findings are not driven by these coding choices.
16. Shaun M. Dynarski, Steven W. Hemelt, and Joshua M. Hyman, "The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes," *Educational Evaluation and Policy Analysis* 37, 1S (2015): 53S–79S.
17. The finding remained consistent across multiple model specifications (Technical Supplement Table S.3), estimation methods (Table S.4), and sample definitions (Table S.6).
18. Bailey, Duncan, Odgers, and Yu (2017); Bailey, Duncan, Cunha, Foorman, and Yeager (2020).
19. The NYC School Survey is administered annually by New York City Public Schools to students, teachers, and parents. The survey was previously titled the "Learning Environment Survey" until 2015. For detailed information on the domains and scoring methodology, see the *NYC School Survey Framework & Scoring Technical Guide* (New York City Public Schools, published annually). Technical Supplement Table S.7 lists the survey items included in each domain's constructs and subconstructs. To assess the internal consistency of the composite constructs used in the analysis, the team computed Cronbach's alpha for each subconstruct and domain within each year, using school-level means as the unit of analysis. Because item availability varies across years, Cronbach's alpha was computed using only the items with non-missing data in each year, and the number of contributing items was recorded alongside each alpha value. Technical Supplement Table S.8 reports the range of alpha values and the range of contributing items across years for each construct. Most subconstruct-level alphas exceed 0.80, and domain-level alphas range from 0.82 to 0.99, indicating considerable internal consistency. Two subconstructs with limited item availability (college encouragement and college application support) have lower alphas in early years when only two items were available but improve substantially in later years as additional items became available.
20. The School Quality Report (SQR) is published annually and provides summary statistics on school demographics, student outcomes, and accountability metrics for all New York City public schools. Data were drawn from the SQR Citywide Results files covering the 2014-2015 through 2023-2024 academic years, downloaded from the New York City Public Schools InfoHub website.
21. See Technical Supplement Table S.10 for a complete list of these measures. In particular, students are considered overage and under-credited when they are older than the typical age for their grade and have not accumulated enough credits to be on track for graduation, a common indicator used in New York City schools to identify students at an elevated risk of dropping out.
22. See Table 1 in Zhu, Martinez, Commins, Solís-Román, and Nash (2026).

23. The reduction in special education classification was statistically significant when special education status was defined by Individualized Education Program classification, the primary measure used in this study. Under an alternative definition based on ungraded service receipt, which was used in earlier INSIGHTS studies, the effect was smaller and not statistically significant (impact estimate = -2.7 percentage points, p-value = 0.226).
24. Collaborative for Academic, Social, and Emotional Learning, "What Is the CASEL Framework?" (website: https://casel.org/fundamentals-of-sel/what-is-the-casel-framework, accessed on May 17, 2026).
25. Allison M. Ryan and Helen Patrick, "The Classroom Social Environment and Changes in Adolescents' Motivation and Engagement During Middle School," *American Educational Research Journal* 38, 2 (2001): 437–460; Aprile Benner, "The Transition to High School: Current Knowledge, Future Directions," *Educational Psychology Review* 23, 3 (2011): 299–328.

Acknowledgments

The authors thank Meghan McCormick, Erin O'Connor, Elise Cappella, and Sandee McClowry, whose foundational work and guidance regarding the original randomized controlled trial and previous INSIGHTS research were instrumental to this study. We are also grateful to James Kemple and Janice Lee at the Research Alliance for New York City Schools for their assistance and support in assembling the data and constructing the analysis file for this long-term follow-up study. We would like to thank Elizabeth Ederer for her irreplaceable role as the project's resource manager and Madelyne Lynam for her incredible work as report coordinator. We are also appreciative of Luisa LaFleur and Carolyn Thomas for their careful editing and preparation of the brief for publication. The study benefited from input and discussions with MDRC colleagues, including William Corrin, Susan Sepanik, and Rebecca Unterman. We also acknowledge Michelle Dixon and Samuel Maves, whose early contributions established the analytic foundation for this work.

The research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A220174 to MDRC. The opinions expressed are those of the authors and do not necessarily reflect the views of the Institute or the U.S. Department of Education.

The following organizations support dissemination of MDRC publications and our efforts to communicate with policymakers, practitioners, and others: Arnold Ventures, Ascendium Education Group, Yield Giving/MacKenzie Scott, and earnings from the MDRC Endowment. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

The findings and conclusions in this report do not necessarily represent the official positions or policies of the funders.

For information about MDRC and copies of our publications, see our website: www.mdrc.org.

New York
200 Vesey Street, 23rd Flr.
New York, NY 10281
Tel: 212 532 3200

Washington, DC
750 17th Street, NW
Suite 501
Washington, DC 20006

Oakland
475 14th Street, Suite 750
Oakland, CA 94612
Tel: 510 663 6372