MATHEMATICA Policy Research

REPORT

FINAL

Year 2 Report of the Atlanta Public Schools Turnaround Strategy

February 1, 2019

Kristin Hallgren
Naihobe Gonzalez
Kevin Kelly
Alicia Demers
Brian Gill

Submitted to:
Atlanta Public Schools
130 Trinity Avenue, SW
Atlanta, GA 30303
Project Officer: Michael LaMont, Executive Director
Contract Number: 091616-01

Submitted by:
Mathematica Policy Research
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005
Project Director: Kristin Hallgren

Reference Number: 50364.400

CONTENTS

APPENDICES

TABLES

FIGURES

EXECUTIVE SUMMARY

The Turnaround Strategy in Atlanta Public Schools (APS) seeks to transform the district's lowest-performing schools and increase students' achievement. Beginning in the 2016–2017 school year, APS provided additional targeted supports to some schools, called **targeted schools**, and partnered with two providers, Kindezi and Purpose Built Schools (PBS), to operate other schools, called **partnership schools**.

APS engaged Mathematica Policy Research to conduct a three-year evaluation of the implementation and impact of the Strategy. The Year 1 Report (Hallgren et al. 2017) included implementation and impact findings from the first year, focusing on in-depth analysis of High Impact Tutoring, one Strategy component provided to targeted schools, as well as PBS' operation of the first partnership school, Thomasville Heights Elementary.

This is the second report on Year 2 of the Strategy, following an interim report (Hallgren et al. 2018) that presented detailed implementation findings in targeted and partnership schools after the second year of the Strategy (2017–2018). This report analyzes the impact of the Strategy in targeted schools after two years and the effectiveness of two specific intervention components offered in those schools in 2017–2018: (1) math and reading specialists, and (2) Communities in Schools (CIS) case management. It also analyzes the impact of the Kindezi and PBS school partnerships. For each of the analyses, we also provide implementation information as context for the impact findings.

Data for this study include administrative data from APS, including students' demographics, attendance, suspensions, and test scores. APS also provided data on which students were on math and reading specialists' rosters, and on CIS site coordinators' caseloads. For the implementation analysis, we conducted site visits in spring 2018 to each targeted and partnership school to interview principals, teachers, and other school staff.

KEY IMPACT FINDINGS—YEAR 2

- As a whole, Turnaround Strategy efforts are producing improvements in math performance in targeted schools. We did not find evidence of positive impacts of schoolwide targeted support on other student outcomes.
- There is little evidence that support from math and reading specialists—one component of the Turnaround Strategy in targeted schools—had an impact on academic outcomes.
- We found no evidence that the CIS case management component of the Turnaround Strategy improved student suspensions, attendance, or academic achievement.
- Turnaround partnership schools are producing improvements in math performance. However, other effects were mixed, varying by outcome and by partner organization.

Discussion and recommendations

DISCUSSION

- The Turnaround Strategy is producing improvements in math performance in both targeted and partnership schools.
- The limited impacts on other outcomes reflect the challenges of successfully improving schools and are consistent with school turnaround literature.
- Specialists did not impact the academic performance of students on their rosters but took on additional responsibilities in the schools that may have provided other value.
- The limited impacts of CIS case management services are consistent with other recent studies, and may be partly explained by implementation challenges.

The Turnaround Strategy is producing improvements in math performance in both targeted and partnership schools. Targeted supports, the Kindezi partnership, and the PBS partnership all improved students' math scores on the statewide Georgia Milestones exams. The analyses capture the effect of the overall set of supports provided to the schools rather than individual components, such as support from math and reading specialists. Because we did not find positive impacts from Strategy components, it is likely that schoolwide changes are driving the positive math impacts. Although we cannot determine with certainty what has driven these improvements, they are a promising sign that the Strategy is supporting growth in students' academic performance in ways that would not be achieved without it.

The limited impacts on other outcomes reflect the challenges of successfully improving schools and are consistent with school turnaround literature. We did not find evidence of positive impacts of schoolwide targeted support or school partnerships on other student outcomes, aside from the Kindezi partnership's impact on English language arts (ELA) achievement. Our findings are consistent with other studies of school turnaround efforts, which have often found limited impacts (for example, Dragoset et al. 2017; Fryer 2014; Research for Action 2018; Zimmer et al. 2017). Continued research about the complexity and nuances of improvement efforts will be helpful for developing a deeper understanding of effective ways to improve schools in need of support.

Specialists did not impact the academic performance of students on their rosters but took on additional responsibilities in the schools that may have provided other value. We found limited evidence that math and reading specialists impacted Georgia Milestones or STAR scores for students on their rosters. Despite these results, implementation findings suggest that staff in targeted schools saw the specialists as highly valuable additions to their school who supported students in ways that could have improved overall learning in the school. In addition, some students who worked with specialists were not on the rosters, which suggests that the analysis may have underestimated the true impact of the specialists. We also found some evidence that math and reading specialists may have been effective with the lowest-performing students they were originally intended to serve. A renewed focus on those students might produce more robust results.

The limited impacts of CIS case management services are consistent with other recent studies, and may be partly explained by implementation challenges. Receiving small group or individual support from a CIS site coordinator did not significantly impact students' likelihood of being chronically absent, likelihood of being suspended, or performance on Georgia Milestones exams. This limited success may have been due to implementation challenges, such as most schools only having one CIS site coordinator at their campus two days a week, high turnover among CIS staff, or the low intensity of small group and individual support that most students experienced. These results are in line with other recent research on CIS case management services (Parise et al. 2017). It is possible that the addition of a CIS site coordinator in targeted schools allowed other nonacademic support staff, such as clinical psychologists and social workers, to provide more support to other students in the schools than would have been possible if the CIS site coordinator had not been present.

RECOMMENDATIONS

- APS should continue to explore how to support students' growth in subjects other than math.
- The district could capture richer program data to better understand which supports are most effective, and for whom.
- The district should monitor the use of suspensions, particularly in partnership schools.
- Extending the supports in the Turnaround Strategy may be critical as schools begin to improve.

APS should continue to explore how to support students' growth in subjects other than math. Aside from the Kindezi partnership at Gideons Elementary, the Strategy has not yet shown improvements in students' ELA achievement. Kindezi strongly emphasized and resourced remediating students' foundational reading skills. Although both targeted and PBS schools have additional staff to support ELA instruction, an even greater focus on remediating foundational reading skills may be necessary. There is also indication that schools should pay additional attention to science and social studies instruction. The Kindezi partnership significantly worsened students' science and social studies scores, and science and social studies performance also declined in targeted and PBS schools, although those impacts were not statistically significant. This finding suggests that there may be a trade-off when improving test scores across different subjects.

The district could capture richer program data to better understand which supports are most effective, and for whom. Capturing the frequency and duration that math and reading specialists meet with students on their rosters and regularly updating those rosters would position the district to better understand the specialists' effectiveness. Similarly, tracking which students work with nonacademic support staff could enable the district to assess the effectiveness of the different types of nonacademic supports available and identify which students could benefit most from each type of support.

The district should monitor the use of suspensions, particularly in partnership schools. The reported number of student suspensions increased in both the PBS and Kindezi partnership schools. This finding should be interpreted cautiously because APS and partnership staff reported that suspensions may have been underreported before the partnerships began, in which case the apparent increase may not be real. Staff from each partnership reported improvements in student behavior but recognized that behavior continued to be a challenge in the schools. In addition, although we did not find that targeted supports led to increased suspensions, staff in those schools also reported behavior issues. Despite efforts to address behavior challenges, suspension rates in partnership and targeted schools remained relatively consistent over the course of the 2017–2018 school year. APS and the partnership organizations should closely monitor suspension rates to assess whether any improvements become evident.

Extending the supports in the Turnaround Strategy may be critical as schools begin to improve. Staff from targeted and partnership schools perceived improvements in students' academics and nonacademics, yet they also pointed to the need to see additional and continued improvements in both areas, especially for students who are still not performing at grade level. School staff also noted that students showed growth in their social and emotional skills but continued to behave in ways that disrupted student learning during class time. APS may want to consider ways of offering academic and nonacademic supports to even more students at low-performing schools. In addition, our findings and other turnaround literature highlight how turning around low-performing schools can take several years. As schools begin to show improvements, APS should plan how to scaffold or extend the Turnaround Strategy supports so schools can effect lasting change.

Next steps for the evaluation

The final year of the evaluation will include an implementation study of the third year of Turnaround Strategy supports as well as additional impact analyses. The evaluation team will collaborate with APS to determine how to focus our analyses in the study's final year. For example, we may examine critical factors of turnaround success, such as school culture, leadership abilities, or staffing mobility and quality (see Center on School Turnaround 2017). The information gained from the study should help APS and other school districts learn more about ways to effectively support low-performing schools.

I. INTRODUCTION

The Atlanta Public Schools (APS) Turnaround Strategy seeks to transform the district's lowest-performing schools and increase student achievement. To promote the transformation of these schools, the Strategy provides three levels of supports: foundational supports for 27 schools, more intensive supports for 6 schools that show greater needs, and additional targeted supports for 13 schools demonstrating the highest needs. These 13 schools, called **targeted schools**, receive resources for implementing academic and nonacademic supports. In addition to schools that receive foundational, intensive, and targeted supports, the Strategy includes schools whose daily operations are overseen and managed by two partner organizations: Kindezi and Purpose Built Schools (PBS). These schools are called **partnership schools**.

APS contracted with Mathematica Policy Research to better understand how schools implement the Turnaround Strategy and its effects on students' academic and behavioral outcomes. Mathematica's evaluation team submitted the first annual report to APS in October 2017. It included implementation findings for the first year of the Strategy (2016–2017) and results from impact analyses of two of the Strategy's components: High Impact Tutoring and the PBS partnership with Thomasville Heights Elementary School.

This is the second report on Year 2 of the Strategy, following an interim report (Hallgren et al. 2018) that presented detailed implementation findings in targeted and partnership schools after the second year of the Strategy (2017–2018). It includes analyses of the impacts of the overall Strategy and its individual components on student outcomes. Chapter II describes the data sources and methodology used to estimate impacts. Chapter III presents the overall impacts of targeted supports after one and two years. Chapters IV and V then discuss the specific impacts of two of the targeted supports: math and reading specialists, and Communities in Schools (CIS) case management services, respectively. Chapter VI presents the impacts of the Kindezi and PBS partnerships in four partnership schools. Chapter VII concludes the report with a discussion of key findings and describes next steps for the evaluation.

KEY IMPACT FINDINGS—YEAR 2

- As a whole, Turnaround Strategy efforts are producing improvements in math performance in targeted schools. We did not find evidence of positive impacts of schoolwide targeted support on other student outcomes.
- There is little evidence that support from math and reading specialists—one component of the Turnaround Strategy in targeted schools—had an impact on academic outcomes.
- We found no evidence that the CIS case management component of the Turnaround Strategy improved student suspensions, attendance, or academic achievement.
- Turnaround partnership schools are producing improvements in math performance. However, other effects were mixed, varying by outcome and by partner organization.

II. DATA SOURCES AND METHODOLOGY

This chapter includes a description of the data and methods used to evaluate the overall Turnaround Strategy in targeted and partnership schools, and two intervention components of the Strategy—math and reading specialists, and CIS case management.

A. Data sources

Administrative data. APS provided several types of administrative data for this study, including student demographic, test score, attendance, and suspension data. Test score data were provided from two assessments: the statewide Georgia Milestones exam, taken by students in grades 3–5, and the districtwide STAR exam, taken by students in all elementary grades. Attendance data are reported as the total number of days a student was absent for each school in which they were enrolled. Students are assumed to be in attendance each day unless a staff member records an absence. Suspension data include the number of days a student was suspended for each disorderly conduct event, as entered by school staff.[1] Appendix A provides additional information on the administrative data used in the study.

Math and reading specialist rosters. APS provided rosters of students who worked with math and/or reading specialists as part of the Strategy. For the purpose of the impact analysis, we defined math and reading specialist students as those who appear on the rosters and worked with a specialist for at least 10 weeks. Apart from the duration of services, no data were available on the frequency or intensity of services.

CIS service data. APS provided a list of students who received case management services from CIS site coordinators. The data included each activity that CIS site coordinators logged with a student and whether it was a whole-school, small group, or individual activity. For the purpose of the impact analysis, we defined CIS case management students as those who had at least five small group or individual activities logged by a CIS site coordinator.

Implementation data. For the implementation analysis, we conducted site visits in spring 2018 to the 13 targeted schools, the Kindezi partnership school, and the 3 PBS partnership schools. Site visits included semi-structured interviews and focus groups with principals, teachers, and other school staff, focusing on respondents' experiences with academic, nonacademic, instructional, and leadership supports. Additional information on implementation data collection, analysis, and findings can be found in the Year 2 interim report (Hallgren et al. 2018).

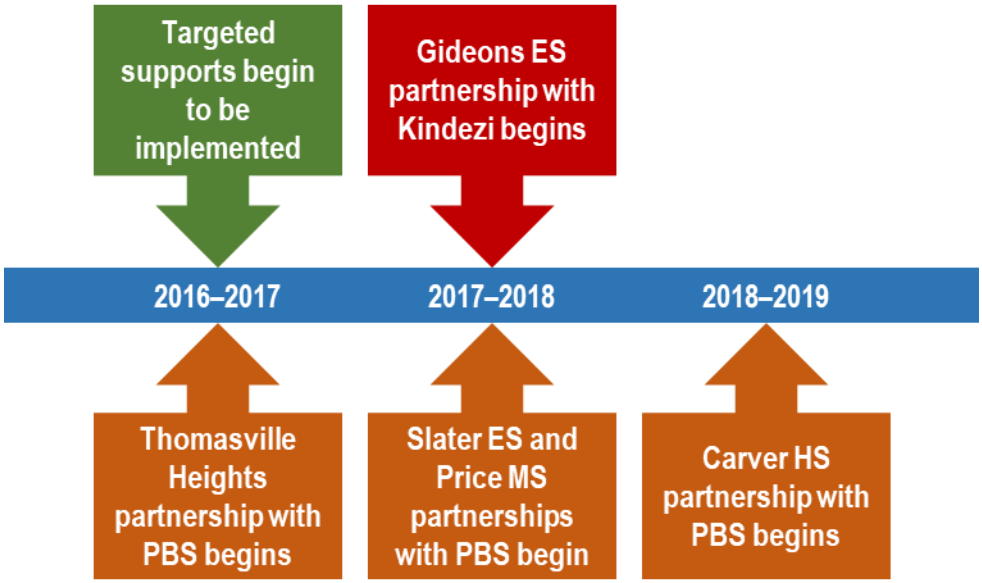
B. Methodology to evaluate overall targeted and partnership school effects

To examine the overall effects of targeted supports and school partnerships, we used a quasi-experimental research design known as difference-in-differences. This design compares changes in outcomes between students at targeted or partnership schools and those in comparison schools before and after APS implemented the Turnaround Strategy. Figure II.1 shows a timeline of when schools first implemented supports from the Strategy. Thirteen schools began receiving

[1] Suspensions include both in school and out of school suspensions.

targeted supports in the 2016–2017 school year, whereas school partnerships were staggered. PBS began operating one school—Thomasville Heights Elementary—in 2016–2017. In the following school year, PBS began operating two additional schools—Slater Elementary and Price Middle School—and Kindezi began operating Gideons Elementary. PBS also began operating Carver High School in the 2018–19 school year. Estimates of PBS' impact at Carver High School will be included in the Year 3 report.

Figure II.1. Timeline of Turnaround Strategy implementation in targeted and partnership schools



To understand schools' trends over time, we analyzed administrative data from the 2012–2013 through 2017–2018 school years. The first difference in the difference-in-differences design compares student outcomes in targeted or partnership schools after the implementation of the Strategy relative to previous years; the second difference compares the targeted or partnership schools' changes to those in comparison schools. The difference-in-differences approach thus accounts for the targeted or partnership schools' prior performance and for stable differences between the targeted or partnership and comparison schools.[2] As a result, comparison schools can be higher (or lower) performing than targeted or partnership schools, but should have had similar trends before APS implemented the Turnaround Strategy.

[2] The difference-in-difference design accounts for differences between the schools, assuming that those differences were consistent over time. Another condition for the difference-in-differences design is that the composition of students at targeted and partnership schools should be stable over the analysis period. If low-performing students left the school after APS introduced the targeted or partnership supports or new high-performing students joined the school since APS introduced the supports, improvements in the school's academic performance could be due to compositional changes rather than the targeted supports or partnerships. In addition, the anticipation of the targeted supports or partnerships should not have caused performance to drop the year before the supports were introduced (for example, due to staff turnover). Otherwise, improvements since the schools became targeted or partnership schools could merely reflect a return to stable performance rather than the impacts of the supports themselves. We do not find evidence of either occurrence. See Appendix B for additional details.

In selecting comparison schools, we considered only APS (non-charter) elementary and middle schools that did not participate in the Turnaround Strategy. In addition, we excluded schools in the North Atlanta and Grady clusters from the comparison group because the demographic profile and average academic performance of these schools, located in the northern and northeastern parts of the city, differed substantially from those that participated in the Strategy. After these exclusions, we identified 15 comparison schools that operated during the six most recent school years in the analysis.[3] These schools were relatively low performing compared to the rest of the district and had trends similar to the targeted and partnership schools before implementation of the Turnaround Strategy.

By comparing targeted or partnership schools to these non-Turnaround Strategy schools, we measured the overall effect of the targeted supports or school partnerships relative to what might have occurred at these schools without the Strategy. Although the comparison schools also may have received other types of support in recent years, those supports likely reflect what targeted and partnership schools would have received if the Strategy had not been implemented. As a result, these analyses measure the impacts of the targeted supports and school partnerships compared to other types of supports low-performing schools in APS might have received.

We examined impacts on six outcomes: student achievement on the Georgia Milestones English language arts (ELA), math, science, and social studies exams; students' likelihood of being suspended; and their likelihood of being chronically absent.[4] To assess which changes represent statistically significant impacts—in other words, impacts that we can confidently conclude are real—we used student-level data to conduct regression analyses that accounted for students' and schools' characteristics.[5] Chapter III presents results for the overall impact of targeted supports; Chapter VI presents impacts of the Kindezi and PBS Turnaround Strategy partnerships.

C. Methodology to evaluate impacts of targeted school components

We estimated the impact of two intervention components used in targeted schools as a part of the Strategy: (1) working with a math or reading specialist during the 2017–2018 school year, and (2) receiving case management support from CIS site coordinators. Each component was offered to a subset of students in targeted schools, so we used a different methodology for these analyses than when estimating the overall impact of targeted and partnership schools.

To measure the effect of each component, we identified a comparison group of students in targeted schools similar to those who received the support but who did not receive the support themselves. In other words, we compared students who worked with a specialist with similar students in targeted schools who did not work with a specialist in the same subject, and students

[3] See Table B.1 in Appendix B for a list of targeted, partnership, and comparison schools.

[4] Students in grades 3 through 5 took the Georgia Milestones exams in ELA and math, but only students in grade 5 took the science and social studies exams. Data on absences and suspensions were available for all grades.

[5] The statistical model also included grade, school, and year fixed effects. Standard errors were clustered at the school level to account for the lack of independence in student outcomes within schools. Additional details on the statistical approach to measuring impacts are available in Appendix B.

who received CIS case management support with similar students in targeted schools who did not. To identify similar comparison students, we used an approach known as propensity-score matching. Following this method, we matched each student who received a support with students in the same grade who had similar prior test scores, attendance, and suspensions, and similar demographic characteristics, but did not receive the support. Propensity-score matching is a well-established approach and has been found to approximate the results of "gold standard" experimental methods (Tuttle et al. 2013; Gill et al. 2015).

Because students in targeted schools received several academic and nonacademic interventions in the 2017–2018 school year as part of the Turnaround Strategy, we selected comparison students from within the same set of targeted schools. Having all of the students in the analyses attending the same schools ensured that matched comparison students had access to the same other services as those who received math or reading specialist support or CIS case management services, and that they shared a similar school environment.

We successfully matched 318 of 458 students who worked with a math specialist with 1,737 comparison students, 300 of 459 who worked with a reading specialist with 1,610 comparison students, and 292 of 348 who worked with a CIS site coordinator with 1,225 comparison students.[6] For each analysis, we confirmed that the matched comparison group had no significant differences with the group that received the support in any of the baseline measures available (see Table C.4 in Appendix C and Table D.4 in Appendix D). However, as with any nonexperimental evaluation, it is possible that differences not captured in the available administrative data could exist.

We examined the impact of each support on student achievement on the Georgia Milestones and STAR exams in math and ELA. Because CIS case management is primarily a nonacademic intervention, we also analyzed its impact on the likelihoods of students being suspended and chronically absent. After matching, we measured impacts on the key outcomes of interest using regression analyses that controlled for small remaining differences in baseline student characteristics. Chapter IV presents the results of the math and reading specialists analysis; Chapter V presents the results for CIS case management services. Additional details about the analytical methods and results of these analyses appear in Appendices C and D.

[6] Across all grades, we could not match 79 math specialist students, 101 reading specialist students, and 38 CIS case management students because they did not have all the baseline data required to conduct the matches. In addition, we did not match a small number of the remaining students because no comparison students resembled them sufficiently.

III. OVERALL FINDINGS FROM TARGETED SCHOOLS

Starting in the 2016–2017 school year, APS provided targeted schools with a variety of leadership, instructional, academic, and nonacademic supports. In response to feedback from school staff, the district slightly modified its implementation of the Turnaround Strategy for the 2017–2018 school year (Figure III.1). Targeted schools continued to receive some supports, such as professional learning, CIS staff, and student support staff.[7] The district also allowed each targeted school to select the academic supports[8] they wanted to implement in their schools from eight options: (1) math and reading specialists, (2) instructional coaches, (3) paraprofessionals, (4) High Impact Tutoring, (5) an intervention block, (6) Saturday school, (7) an extended school day, and (8) Spring Break Academy. All schools had at least one math specialist, reading specialist, and instructional coach; most schools brought in additional staff for each role. A few schools added paraprofessionals, tutoring, intervention blocks, Saturday school, or an extended school day in their schools, but none chose to offer the Spring Break Academy. The Year 2 interim report provides additional information on each of the targeted supports (Hallgren et al. 2018).

Figure III.1. Key components offered to Turnaround Strategy targeted schools in the 2017–2018 school year

*This support was newly offered to targeted schools in the 2017–2018 school year.

[7] All targeted schools had one student support staff, such as a social worker or counselor, and one CIS site coordinator. Hollis Innovation Academy had two CIS site coordinators in the 2017–2018 school year, but only one was funded through the Turnaround Strategy.

[8] In the second year of the Strategy, principals selected academic supports according to a fixed budget provided by APS. For example, one school discontinued High Impact Tutoring, which it used during the first year of the Strategy, and added another math specialist, a reading specialist, and the intervention block in the second year (in addition to retaining math and reading specialists, two coaches, and the student support practitioner). Another school discontinued High Impact Tutoring and added paraprofessional support and an instructional coach (in addition to retaining math and reading specialists, two coaches, and the student support practitioner).

A. Impact of targeted supports

KEY IMPACT FINDING: As a whole, Turnaround Strategy efforts are producing improvements in math performance in targeted schools. We did not find evidence of positive impacts of schoolwide targeted support on other student outcomes.

As described in Chapter II, the difference-in-difference research design used in this analysis measures the overall impact of targeted supports by assessing changes in schoolwide student achievement relative to pre-intervention achievement in targeted schools and the trends of comparison schools. In Figure III.2, we graph standardized mean student performance on the state exams[9] in ELA, math, science, and social studies, as well as suspension and chronic absenteeism rates, for the targeted and comparison schools. These figures show that the comparison schools had similar trends before the 2016–2017 school year, even though they had higher average student outcomes.

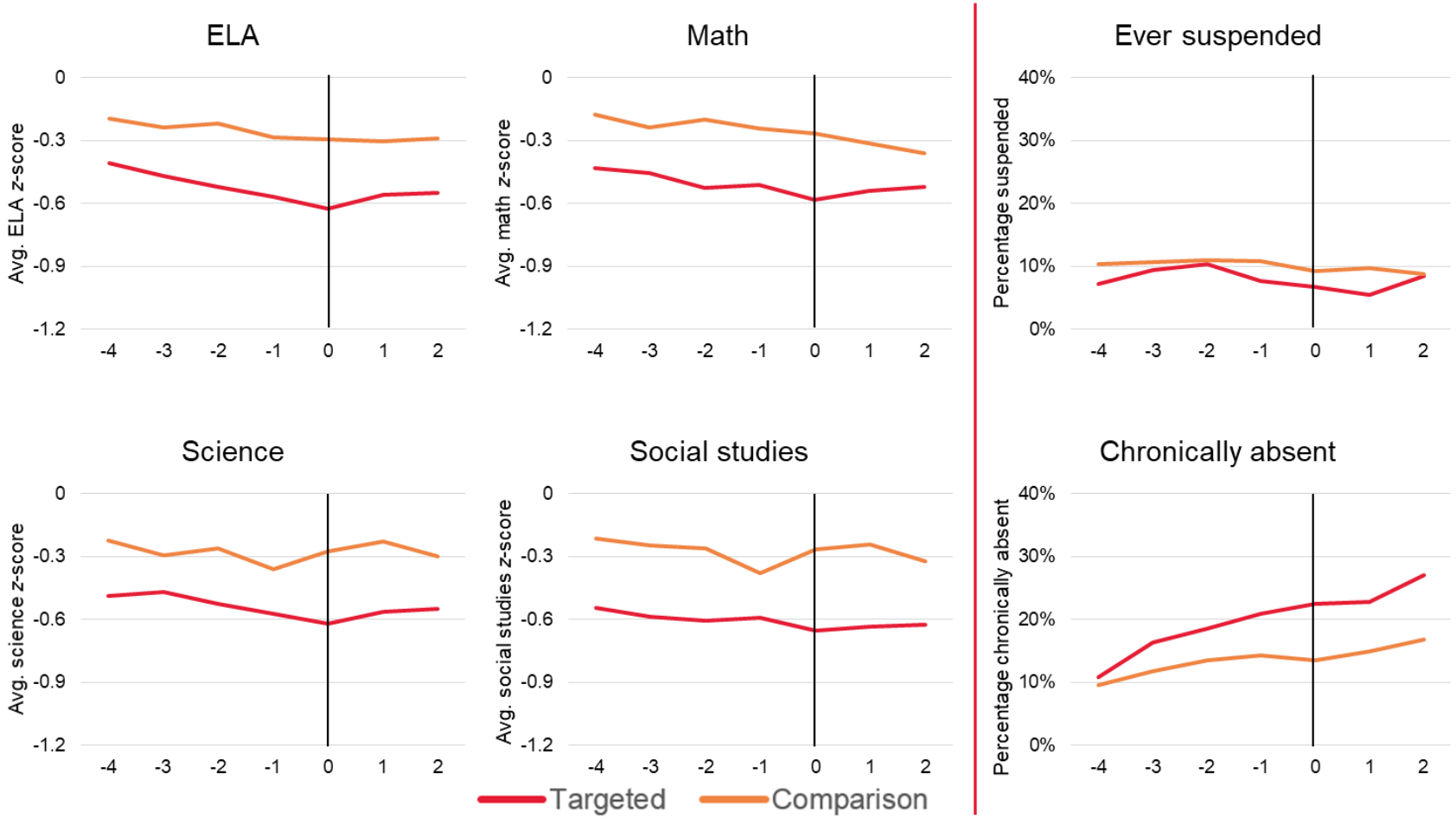
If the targeted supports were effective, we would expect the differences between targeted and comparison schools to shrink after the intervention. The figures show a few small changes in targeted schools relative to comparison schools after APS introduced the targeted supports. Most notably, students' math scores increased in targeted schools, whereas those in comparison schools declined. However, many of the differences in trends were fairly stable even after APS introduced the targeted supports, indicating that it is unlikely the targeted supports affected those outcomes.

The impact analysis provides a more formal test of the significance of the trajectory changes observed in the graphs. The results of this analysis indicate that the second year of targeted supports led to a statistically significant improvement in students' math achievement on the Georgia Milestones (that is, we can confidently conclude that there was an impact).Table III.1 presents the measured impacts, where zero represents no impact (that is, students in targeted schools did as well as would have been expected without the additional supports). The impacts on test scores are shown in standardized *z*-score units. The second year of supports provided to targeted schools led to a statistically significant increase of 0.13 standard deviations in math scores. For the typical student in grades 3 to 5, growth of 0.13 standard deviations in math is roughly equivalent to two additional months of learning.[10] Targeted supports had no statistically significant impacts in the first year and no significant impacts on ELA, science, or social studies results, or students' likelihood of being suspended or chronically absent in the second year.

[9] Because scales on state tests changed over time, we converted all scaled scores to normalized *z*-scores that show each student's position in the districtwide distribution. Zero represents the districtwide mean score; positive scores are above the district mean, and negative scores are below it. To make the tested grades consistent over time, we analyzed science and social studies performance in grades 5 and 8 only.

[10] This conversion is based on an analysis of annual learning growth on nationally normed exams (Bloom et al. 2008). To convert impacts into months of learning, we divided the impact estimate by the average of the typical annual growth for students in grades 3 through 5 and assumed a nine-month school year. The accuracy of this conversion depends on the extent to which the learning growth on the Milestones exam is similar to the exams analyzed in Bloom et al. (2008). According to this analysis, annual student growth in math is 0.89 standard deviations for grade 3, 0.52 standard deviations for grade 4, and 0.56 standard deviations for grade 5.

Figure III.2. Changes in outcomes in targeted and comparison schools over time



Source: APS administrative data.

Note: These figures display changes in outcome trends for the Georgia Milestones exams in *z*-scores (standard deviations) and percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the percentage of students ever suspended during the school year. "Chronically absent" refers to the percentage of students missing 10 percent or more of days enrolled.

APS = Atlanta Public Schools; ELA = English language arts.

Table III.1. Impacts of targeted school supports after one and two years

	Academic outcomes (standard deviations)				Nonacademic outcomes (percentage points)	
	ELA	Math	Science	Social studies	Likelihood of being suspended	Likelihood of being chronically absent
Impact on targeted schools in Year 1	-0.01 (0.03)	0.03 (0.04)	-0.11 (0.07)	-0.10 (0.07)	-0.00 (0.02)	0.02 (0.01)
Impact on targeted schools in Year 2	0.01 (0.03)	0.13** (0.04)	-0.05 (0.08)	-0.04 (0.07)	0.02 (0.01)	0.04 (0.02)
Number of students	**22,658**	**22,654**	**12,875**	**12,876**	**40,105**	**40,105**
Number of schools	**28**	**28**	**28**	**28**	**28**	**28**

Source: APS administrative data.

Note: This table displays impact estimates for the 2017 and 2018 Georgia Milestones exams in *z*-scores (standard deviations) and percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the likelihood that a student was ever suspended during the school year. "Chronically absent" refers to the likelihood that a student was missing 10 percent or more of days enrolled. Standard errors are displayed in parentheses below each impact estimate. The sample size reflects the total number of students in each analysis across years.

**Impact is statistically significant at the 1 percent level.

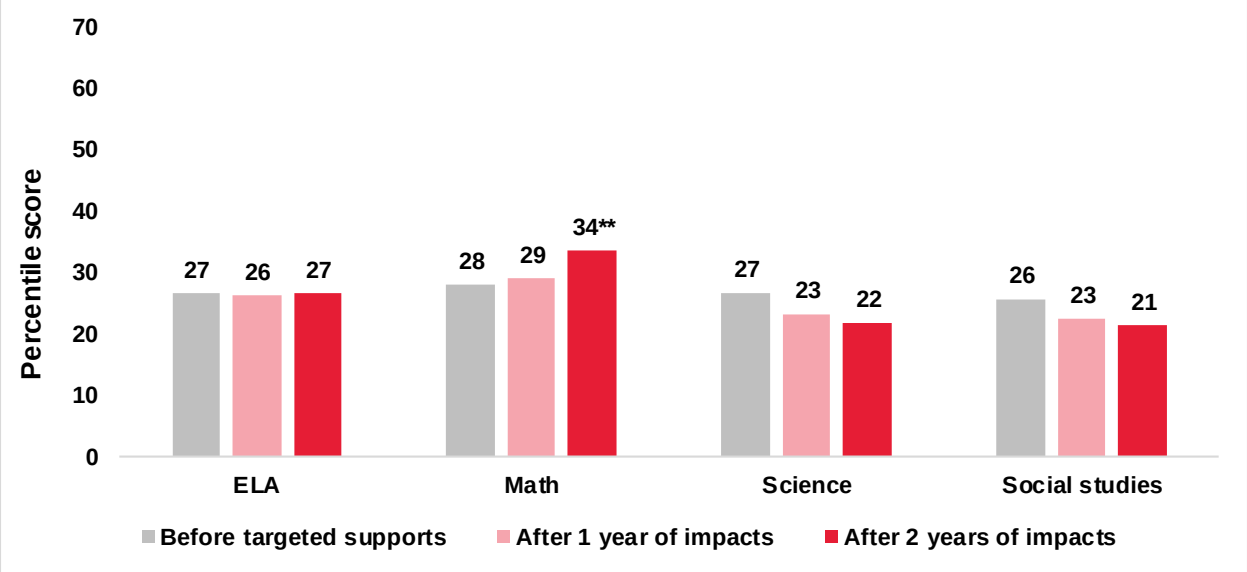
*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; ELA = English language arts.

The impacts on test scores can also be interpreted in terms of changes in students' percentile scores after the targeted schools implemented the Turnaround Strategy. Figure III.3 shows how the average performance of students in targeted schools changed following one or two years of impacts. Academic performance is expressed as the percentile rank of students in targeted schools relative to those in the rest of the district. For example, a student in the 28th percentile performed as well as or better than 28 percent of other students in the district. The gray bars on the left indicate actual average performance in the year before beginning targeted supports. As seen in Table III.1, there was a positive, statistically significant impact on math scores in Year 2. This impact is equivalent to average math performance increasing from the 28th to the 34th percentile. Science and social studies performance decreased, but these impacts were not statistically significant.

Despite the limited impacts of targeted supports overall, the second-year impact on math achievement is a promising sign, as many turnaround efforts fail to produce any measureable positive impacts in a comparable or even longer period (for example, Dragoset et al. 2017; Gill et al. 2007; Dougherty and Weiner 2017; Heissel and Ladd 2018). When impacts are detected, positive impacts tend to be larger in math than reading, so those measured for targeted school supports follow the trends of turnaround efforts elsewhere (Fryer 2014; Zimmer et al. 2017).

Figure III.3. Impacts of targeted supports on academic outcomes one and two years after introducing supports



Source: APS Administrative data.

Notes: We calculated the percentile ranks after one and two years of targeted supports by adding the *z*-score impact estimates (Table III.1) to the average *z*-scores of targeted school students on the Georgia Milestones exams in the 2015–2016 school year. We assumed that the percentile rank of the average comparison student did not change between years.

**Impact is statistically significant at the 1 percent level.

ELA = English language arts.

B. Implementation of targeted supports

IMPLEMENTATION CONTEXT

- More than three-quarters of the targeted schools continued to face pervasive behavioral issues and wanted more nonacademic supports for all students.
- Teachers faced challenges in improving academic achievement because of limited capacity, "compassion fatigue," or burnout.
- When planning for next year, school leaders from targeted schools want more of the existing nonacademic supports but feel compelled to add academic supports because academic outcomes can be measured more easily.

Findings from the Year 2 implementation study shed light on the challenges that staff in targeted schools experienced when working to improve their schools and may help explain the limited impacts of targeted supports in the first two years of the Strategy on outcomes other than math. Below we discuss key findings that provide context for the results of the impact analyses.

More than three-quarters of the targeted schools continued to face pervasive behavioral issues and wanted more nonacademic supports for all students. Staff from all schools reported noticing positive changes in student behavior that they attributed to nonacademic

supports. However, school leaders and staff from at least nine targeted schools said that students continued to demonstrate moderate to severe behavioral issues that regularly interrupted instruction or required additional staff support. One indicator of behavior issues is school suspensions (although the accuracy of this indicator depends on suspension recording practices). According to APS data, in the 2017–2018 school year, 9 percent of students in targeted schools were suspended, compared to 5 percent in the 2016–2017 school year (Figure III.2). It is therefore possible that students' behavior remained the same (or worsened) despite the additional supports provided to targeted schools. In light of these heightened behavioral issues, staff emphasized the need for nonacademic supports for all students at their school. Insufficient nonacademic supports to meet the needs of all students may have prevented measurable improvements in student outcomes.

Teachers faced challenges in improving academic achievement because of limited capacity, "compassion fatigue," or burnout. Leaders and staff in about half of the schools discussed the need for teachers to improve their capacity or skills. School leaders stated that teachers needed additional improvements in their instructional practices, deeper content knowledge, or a quicker response to feedback on their instructional practices. These leaders emphasized the importance of teachers rapidly adjusting to support the academic growth of low-performing students. Teachers at these schools explained that teaching the grade-level curriculum to students struggling with foundational skills and frequent, pervasive disruptions to instructional time made growth in academic achievement difficult, especially at a rapid pace.

In three of the schools, school leaders and staff described how teachers struggled to work in a "tough environment." Leaders also noted that teachers often experienced "compassion fatigue" from working with students, parents, and communities that faced chronic challenges, such as homelessness, violence, and poverty. They also noted that burnout rates were high because teachers were working in a climate of constant change, had to take on multiple roles, and felt a great deal of pressure. Teachers explained that they never felt they had enough time for all of their duties because they felt they had more to do than teachers in non-turnaround schools; for instance, they described needing more time to plan or modify their instruction to accommodate students who were not at grade-level.

When planning for next year, school leaders from targeted schools want more of the existing nonacademic supports but feel compelled to add academic supports because academic outcomes can be measured more easily. Eight principals described the need for additional nonacademic supports to address different types of issues. For example, staff at one school reported that they would benefit from an additional clinical specialist and CIS coordinator. Staff at two schools said that part of addressing students' needs included receiving more intensive resources, such as referrals to health care specialists. One school also described how it used external grants to fund resources, such as an in-school clinic that provided physical and mental health services to students and their families.

Even with this need for nonacademic supports, at least two principals reported choosing additional academic supports for the coming year instead of nonacademic support staff. They gave two reasons for this decision. First, they wanted to build on the academic achievements that they associated with the academic supports. For example, one principal chose to add another reading specialist over a nonacademic practitioner because the academic growth among students

who worked with specialists seemed greater than that of students working with nonacademic support staff. Second, principals struggled to determine the best way to measure growth for students receiving nonacademic supports. For example, one principal used a mix of academic achievement, attendance, and behavioral data to assess the effectiveness of nonacademic supports but acknowledged that these data may not accurately capture the impact of these supports.

Although staff from targeted schools appreciated the targeted supports available and viewed them as effective for their schools, the current level or type of nonacademic supports may not sufficiently meet the needs of all students. Providing additional resources to address students' nonacademic needs may allow schools to support more students and help increase teachers' capacity to drive academic achievement. In response to feedback from the targeted schools in 2017–2018, APS provided all targeted schools with full-time CIS coordinators in 2018–2019. APS also continued to offer a menu of academic support options to principals for the 2018–2019 school year.

Next, we turn to examining the effectiveness of two of the components provided to targeted schools as a part of the Turnaround Strategy: math and reading specialists, and CIS case management services.

IV. MATH AND READING SPECIALISTS IN TARGETED SCHOOLS

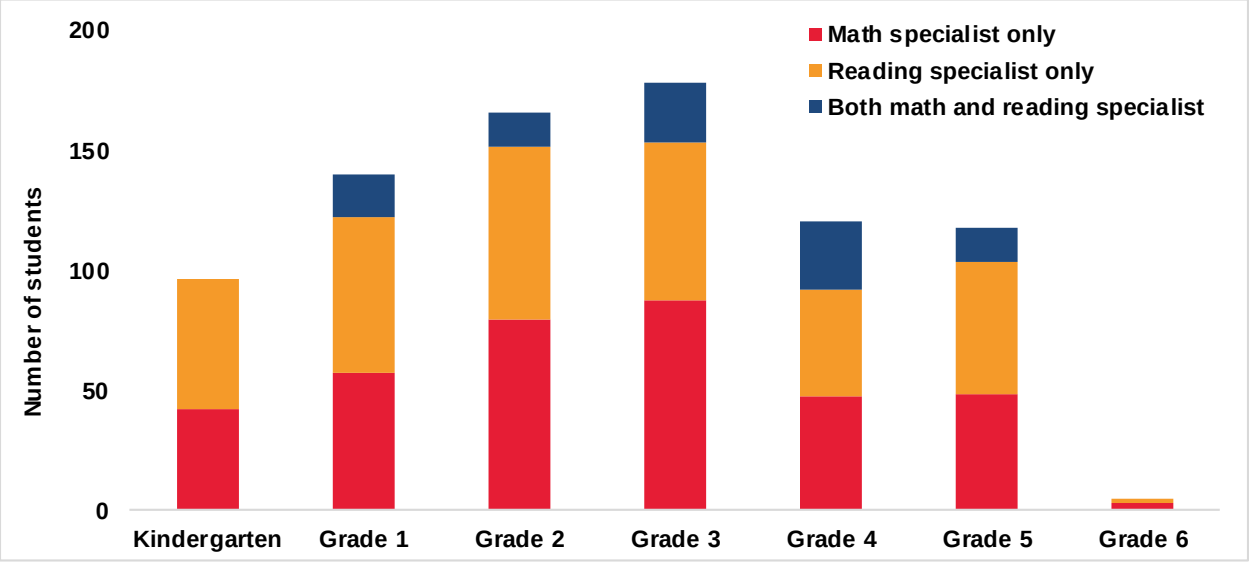
APS provided each targeted school with at least one math specialist and one reading specialist as part of the Turnaround Strategy. This chapter will describe the support that math and reading specialists provided in targeted schools, present the impact that specialists had on the students on their rosters, and discuss key findings from the implementation study, which provide additional context to help interpret the impact results.

A. Description of math and reading specialist support

In the 2017–2018 school year, each targeted school received one math and one reading specialist to help low-performing students build foundational skills. Eleven of the 13 schools also chose to add one or two additional specialists to support more students.[11] Specialists were expected to meet with small groups of the school's lowest-performing students—which APS defined as students in the school's lowest 5 to 10 percent based on performance on student assessments—about three times each week. Specialists used a district-selected curriculum (*Do the Math* for math and *Leveled Literacy Intervention* for reading) to guide small group sessions.

In total, 464 students in targeted schools worked with a math specialist and 460 students worked with a reading specialist, according to roster data. Figure IV.1 shows how many students worked with specialists in each grade. Most schools offered support to students in all grades, but some focused on a subset. The figure also shows how many of the students in each grade worked with only a math specialist, only a reading specialist, or both; relatively few students worked with both.

Figure IV.1. Grade levels of students served by math and reading specialists



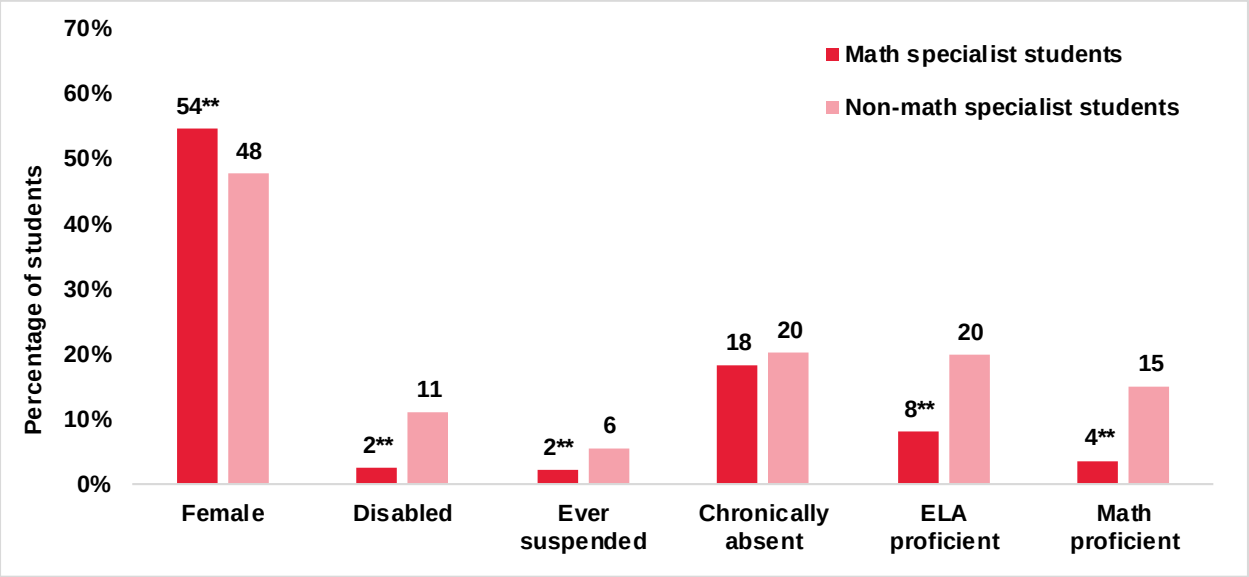
Source: APS administrative data.

APS = Atlanta Public Schools.

[11] Of the 13 targeted schools, 10 had two reading specialists and 3 had one reading specialist; 9 targeted schools had two math specialists and 4 had one math specialist.

Students served by math and reading specialists had below-average prior achievement levels. At the beginning of the 2017–2018 school year, fewer students who worked with math and reading specialists scored "proficient" or "distinguished" on the math and ELA district-wide STAR exams than their classmates in targeted schools who did not work with specialists (Figures IV.2 and IV.3). In addition, fewer of the students who worked with specialists had diagnosed disabilities than did their classmates, and math specialist students were more likely to be female and less likely to have ever been suspended than their classmates.

Figure IV.2. Baseline characteristics of students working with a math specialist compared with students not working with a math specialist in targeted schools



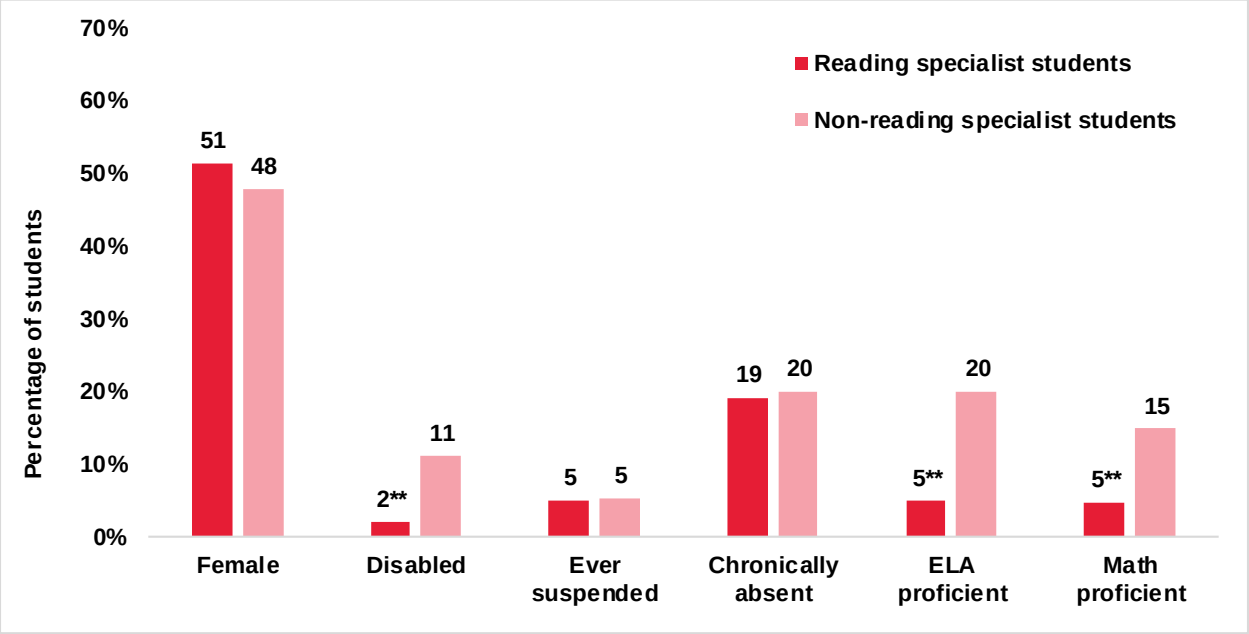
Source: APS administrative data.

Notes: "Ever suspended" shows the percentage of students suspended at least one time in the 2016–2017 school year. "Chronically absent" shows the percentage of students that missed 10 percent or more of days enrolled in the 2016–2017 school year. "ELA proficient" and "Math proficient" show the percent of students scoring at least proficient on the STAR fall 2017 assessment for the respective subject.

**Difference is statistically significant at the 1 percent level.

APS = Atlanta Public Schools; ELA = English language arts.

Figure IV.3. Baseline characteristics of students working with a reading specialist compared with students not working with a reading specialist in targeted schools



Source: APS administrative data.

Notes: "Ever suspended" shows the percentage of students that were suspended at least one time in the 2016–2017 school year. "Chronically absent" shows the percentage of students that missed 10 percent or more of days enrolled in the 2016–2017 school year. "ELA proficient" and "Math proficient" show the percent of students scoring at least proficient on the STAR fall 2017 assessment for the respective subject.

**Difference is statistically significant at the 1 percent level.

APS = Atlanta Public Schools; ELA = English language arts.

B. Impact of math and reading specialist support

KEY IMPACT FINDING: There is little evidence that support from math and reading specialists—one component of the Turnaround Strategy in targeted schools—had an impact on academic outcomes.

To analyze the impact of math and reading specialists on students' assessment scores, we matched students who worked with specialists with similar students in targeted schools who did not receive such support in that subject (see Chapter II for more information on the methods and data used for the analysis). The matched comparison group of students had baseline achievement levels and other characteristics similar to those of students who worked with math and reading specialists (Table C.4 in Appendix C).

In spring 2018, after working with specialists for the year, participating students' achievement levels on Georgia Milestones assessments (administered in grades 3–5) and STAR

assessments[12] (administered in grades 1–5) had marginally improved relative to the comparison students in both math and ELA. These small improvements are not statistically significant; we cannot say with confidence that working with a specialist produced academic gains. Table IV.1 summarizes the impacts of working with a specialist. The results are shown in standardized *z*-score units.

Table IV.1. Impacts of math and reading specialists on Georgia Milestones (grades 3–5) and STAR (grades 1–5) scores

	Math	ELA
Impact of specialists on Georgia Milestones *z*-scores (grades 3–5)	0.06 (0.03)	0.06 (0.04)
Number of students included in analysis	**1,418**	**1,347**
Impact of specialists on STAR *z*-scores (grades 1–5)	0.07 (0.04)	0.05 (0.03)
Number of students included in analysis	**1,702**	**1,636**

Source: APS administrative data.

Notes: This table displays impact estimates in *z*-scores (standard deviations) on the 2018 Georgia Milestones exams and spring 2018 STAR exams for the same subject as the specialist services. Standard errors are displayed in parentheses below each impact estimate. We applied a multiple comparisons adjustment to account for the two exam outcomes in each subject. The sample size reflects the total number of specialist students and matched comparison students in each analysis.

**Impact is statistically significant at the 1 percent level.

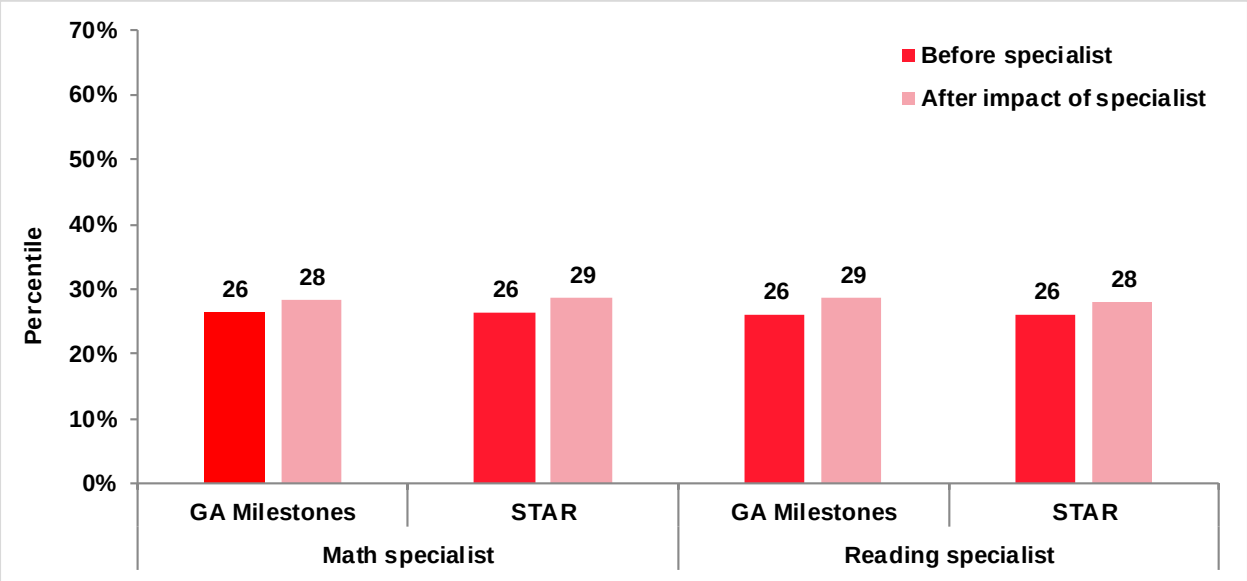
*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; ELA = English language arts.

The results can also be interpreted in terms of how the average student's performance changed after receiving support from the specialists, as expressed by students' percentile rank on assessments (Figure IV.4). If working with a specialist improved academic performance, the lighter bar on the right should be higher than the darker bar on the left for each assessment. As shown in the figure, the average student improved slightly in both the Georgia Milestones and STAR exams for both subjects. However, these improvements were fairly small and, as previously discussed, not statistically significant.

[12] As with the fall 2017 STAR exam, we used ELA scores and then STAR early literacy (SEL) scores if available to examine performance in spring 2018. SEL assessments are most commonly taken by students in early grades in place of STAR ELA assessments. The results were not sensitive to using SEL scores when students did not have ELA scores, but allowed us to include additional students. In the regression, we included an indicator if the SEL exam was used as the outcome. About 2.1 percent of students in the ELA analysis had an SEL assessment score as their outcome.

Figure IV.4. Impact of working with math and reading specialists on student assessments



Source: APS administrative data.

Notes: Percentile ranks before specialists are based on the fall 2017 STAR scores in the same subject in which student received support. We calculated the percentile ranks after the impact of specialists by adding the *z*-score impact estimates to specialist students' average fall 2017 *z*-scores on the STAR exam in the relevant subject. We assumed that the percentile rank of the average comparison student did not change over time.

APS = Atlanta Public Schools; ELA = English language arts.

To better understand the effectiveness of math and reading specialists, we examined whether they had an impact on student performance in schools where students worked with them for a longer period of time (about 7.5 months or more, compared to less time).[13] There is no statistically significant evidence that math or reading specialists improved Georgia Milestones or STAR scores for roster students in these schools (Table C.5 in Appendix C). We also examined whether specialists' impacts varied based on the prior performance level of the students with whom they worked (Table IV.2). There is suggestive evidence that specialists may be effective in improving the achievement of especially low-performing students.[14] Reading specialists significantly improved ELA Georgia Milestones scores for students who were initially the lowest performing in their schools by about 0.12 standard deviations (or almost two months of learning), but did not significantly improve scores for higher performing students. Reading specialists' impacts on STAR scores, and math specialists' impacts on both Georgia Milestones and STAR scores, were also greater among these lower performing students, although they were not statistically significant. Appendix C includes more information on these analyses.

[13] In a quarter of schools, students received reading support for less than 7.4 months and math support for less than 7.7 months on average. This duration was calculated using the start and end dates of support for each student. Most schools did not provide end dates, so we imputed them using the last day each student was enrolled in that school, assuming that specialists provided support through the end of the school year.

[14] Lower-performing students are those whose fall 2017 STAR exam scores were in the lowest 25 percent of their school.

Table IV.2. Impacts of math and reading specialists on assessment scores, by baseline academic performance

	Math		Reading	
	Milestones Math	STAR Math	Milestones ELA	STAR ELA
Impact of specialist support on higher-performing students	0.05 (0.04)	0.05 (0.05)	0.04 (0.05)	0.04 (0.03)
Impact of specialist support on lower-performing students	0.08 (0.05)	0.12 (0.06)	0.12* (0.05)	0.05 (0.04)
Difference in impacts between lower- and higher-performing students	0.03 (0.06)	0.07 (0.08)	0.08 (0.06)	0.01 (0.05)
Number of students	**1,418**	**1,702**	**1,347**	**1,636**

Source: APS administrative data.

Notes: This table displays impact estimates in *z*-scores (standard deviations) on the 2018 Georgia Milestones exams and spring 2018 STAR exams for the same subject as the specialist services. "Lower-performing students" are those who scored in the bottom quartile on the fall 2017 STAR exam of the respective subject compared to other students in their school. "Higher-performing students" are those who scored above the bottom quartile. The Georgia Milestones estimates are based only on students in grades 3 through 5. Standard errors are displayed in parentheses below each impact estimate. We applied a multiple comparisons adjustment to account for the two exam outcomes in each subject. The sample size reflects the total number of specialist students and matched comparison students in each analysis.

**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools.

C. Implementation of math and reading specialist support

IMPLEMENTATION CONTEXT

- Specialists took on many other responsibilities to support non-roster students and teachers.
- Some students who met regularly with specialists may not appear on the rosters.
- Specialists regularly worked with students outside of the bottom 5 to 10 percent of their school because staff felt that more students could benefit from their support.
- Staff saw the math and reading specialists as very effective in supporting academic achievement and building students' foundational skills.

Although the impact findings indicate that math and reading specialists did not improve students' academic achievement, staff in targeted schools saw high value in the addition of specialists to their schools. There are also some indications that specialists' support benefited students in the comparison group, which could limit the impact findings. The implementation findings below showcase ways in which specialists supported academic growth in targeted schools beyond meeting with students on their rosters.

Specialists took on many other responsibilities to support non-roster students and teachers. Specialists and administrators at all targeted schools reported that the specialists took

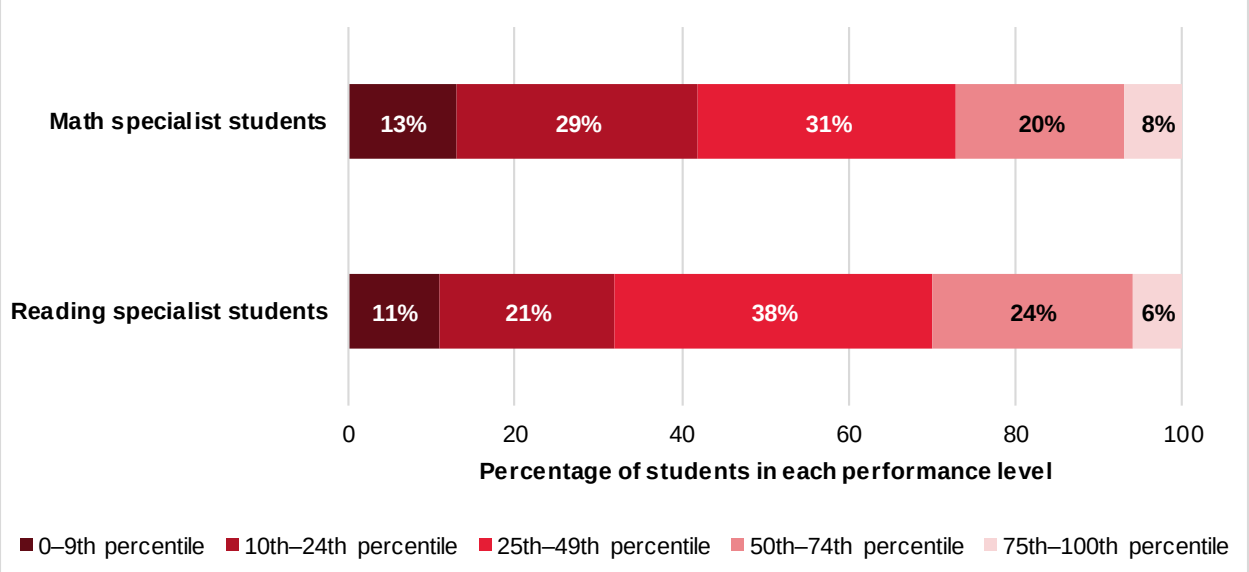
on more responsibilities than just meeting with the students on their rosters. Specialists at 10 schools reported working with additional students outside of their roster groups, including meeting regularly with groups of high-achieving students, pushing into classrooms to work individually or in small groups with students who were struggling to comprehend content, or inviting non-roster students to join select sessions with the students on the roster. Specialists at 8 schools reported supporting teacher professional learning by modeling lessons for teachers, leading professional development sessions, or providing resources used in their small groups to teachers for use in classroom instruction. All of these activities could have improved the achievement of students in the matched comparison group, which would cause the impacts of the specialists to be underestimated.

Some students who met regularly with specialists may not appear on the rosters. Specialists at a few schools noted that they worked with students in grades or subjects outside of the roster data used in the impact analysis. For example, a math specialist at one targeted school reported working with groups of students in grades 2 through 5; however, according to the roster data, there were no math specialist students on the roster at that school. In addition, staff at six schools reported making substantial changes to rosters that were not reflected in the roster data. According to school staff, roster changes were typically made in response to improvements in students' performance after working with the specialist or to replace groups of students in lower-level grades with groups of students in upper-level grades in preparation for the Georgia Milestones. If some students in the matched comparison group actually received support from the specialists, the analysis could underestimate the true impacts of the specialists.

Specialists regularly worked with students outside of the bottom 5 to 10 percent of their school because staff felt that more students could benefit from their support. Although math and reading specialists were expected to work with students in the bottom 5 to 10 percent of academic performance at their school, they worked with students across a variety of achievement levels relative to their classmates (Figure IV.5).[15] At more than half of the schools, staff explained that they believed all students would benefit from working with math and reading specialists because many students lacked the foundational skills that the specialists aimed to build. Specialists reported using STAR and Georgia Milestone assessment data to identify the lowest-performing students in the school. However, they explained that the majority of the school's students did not meet grade-level standards and needed additional support to build their math and reading foundational skills. School leaders explained that they typically assigned students to either a math or a reading specialist—but not both—to maximize the number of students who could be served. In addition, students who were already receiving additional supports, such as special education or English as a Second Language services, were not eligible to work with the specialists.

[15] We also examined the distribution of students on math and reading specialists' rosters by their performance compared to other students in the district (Figure C.1). The majority of students on the rosters were low performing compared to other students in the district, but only 25 percent of math specialist students and 20 percent of reading specialist students were in the lowest 10 percent of students across the district.

Figure IV.5. Baseline performance of math and reading specialist students on the fall 2017 STAR assessment, relative to other students in the same school



Source: APS administrative data.

Notes: Figure reads: Thirty-eight percent of students who worked with reading specialists scored in the 25th to 49th percentile of students in their school on the fall 2017 STAR assessment.

APS = Atlanta Public Schools.

Staff saw the math and reading specialists as very effective in supporting academic achievement and building students' foundational skills. Specialists were largely perceived as being very effective in supporting academic achievement by school staff. In light of the impact findings, staff's perceptions of the effectiveness of the specialists indicate that they may have brought value to the schools beyond what the impact analysis results show. For example, the implementation findings indicate that specialists may have impacted the performance of students not on their roster or improved teachers' instruction. It is also possible that school staff misjudged the effectiveness of the specialists for particular students. For example, supplemental analyses suggest that specialists may be effective in improving the achievement of especially low-performing students—the population that the specialists were originally expected to serve.

If students outside of those listed on the roster regularly worked with math and reading specialists or benefited from their presence through improved classroom teaching, then the actual impact that specialists had may be larger than what our results found. For example, the overall improvements in math performance in targeted schools (discussed in Chapter III) may be due in part to having a math specialist who assisted with improving teachers' math instruction and supported individual students' math comprehension.

As the Turnaround Strategy continues, APS should ensure that rosters for the math and reading specialists are complete and accurate as a way to better assess impacts. In addition, collecting data on how frequently and for how long specialists met with students could be helpful in determining the how the amount of support that specialists provide affects academic performance.

V. COMMUNITIES IN SCHOOLS (CIS) CASE MANAGEMENT IN TARGETED SCHOOLS

APS provided targeted schools with two nonacademic supports in 2017–2018: a full-time student support practitioner[16] (such as a behavioral specialist or social worker) and a half-time CIS site coordinator.[17] Both positions typically worked with a caseload of students identified by the school's Care Team, composed of school staff members who identified students with high nonacademic needs and determined how best to support them. This chapter describes the specific support that CIS site coordinators provided to students on their caseloads; reports the impact they had on caseload students' suspension rates, chronic absenteeism rates, and assessment scores; and discusses key implementation findings to provide additional context for the impact results.

A. Description of CIS case management services

Almost all CIS site coordinators supported two schools. Typically, each coordinator spent two days each week in each of their assigned schools and one day at the CIS central office for planning and professional development. While at a school, site coordinators were to split their time evenly between supporting students on their caseloads and providing whole-school services to support the entire student body. Case management services primarily consisted of individualized wraparound supports, such as checking in on students' attendance, connecting students with tutors, and connecting the students' families to health services. APS anticipated that each site coordinator would have a caseload of 25 to 30 students in each school they served. Whole-school services managed by CIS coordinators, such as a food pantry for students' families or a schoolwide party for perfect attendance, were available to all students in the school.

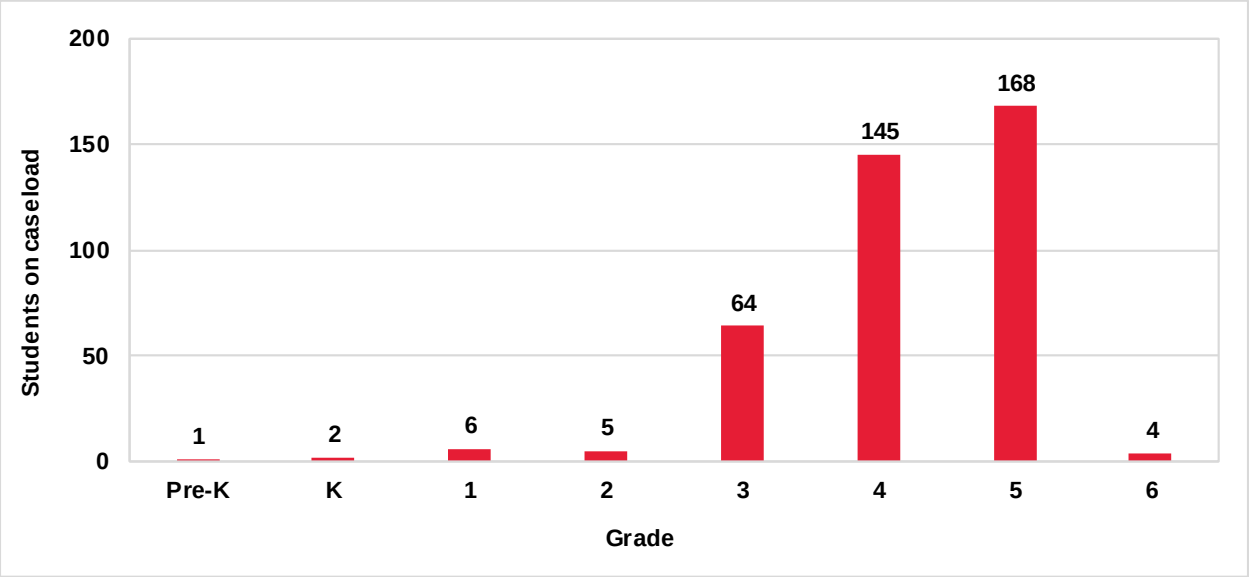
Students on CIS site coordinators' caseloads were almost exclusively in grades 3–5, as recommended by APS (Figure V.1). Within each school, caseloads ranged from 21 to 34 students.[18] Almost half of the students on the site coordinators' caseloads in 2017–2018 had worked with a CIS site coordinator in the 2016–2017 school year. As shown in Figure V.2, CIS caseload students differed in several ways from other students in targeted schools. They were less likely to be male, disabled, or to have been chronically absent in the previous school year. They also had lower baseline math and ELA achievement in fall 2017 and were more likely to have been suspended in the previous school year.

[16] Principals selected which practitioner role—behavioral specialist, clinical therapist, counselor, or social worker—to include at their schools. Depending on the role, practitioners provided support to students on their caseloads for special needs, mental health, trauma, and at-risk behaviors. They often worked with students outside of the classroom and communicated with parents or guardians and other stakeholders as needed.

[17] Barack and Michelle Obama Elementary School had a full-time CIS site coordinator. Hollis Innovation Academy had two CIS site coordinators, one of whom was funded through the Turnaround Strategy.

[18] This range does not include Hollis Innovation Academy, which served 84 caseload students but had two CIS site coordinators.

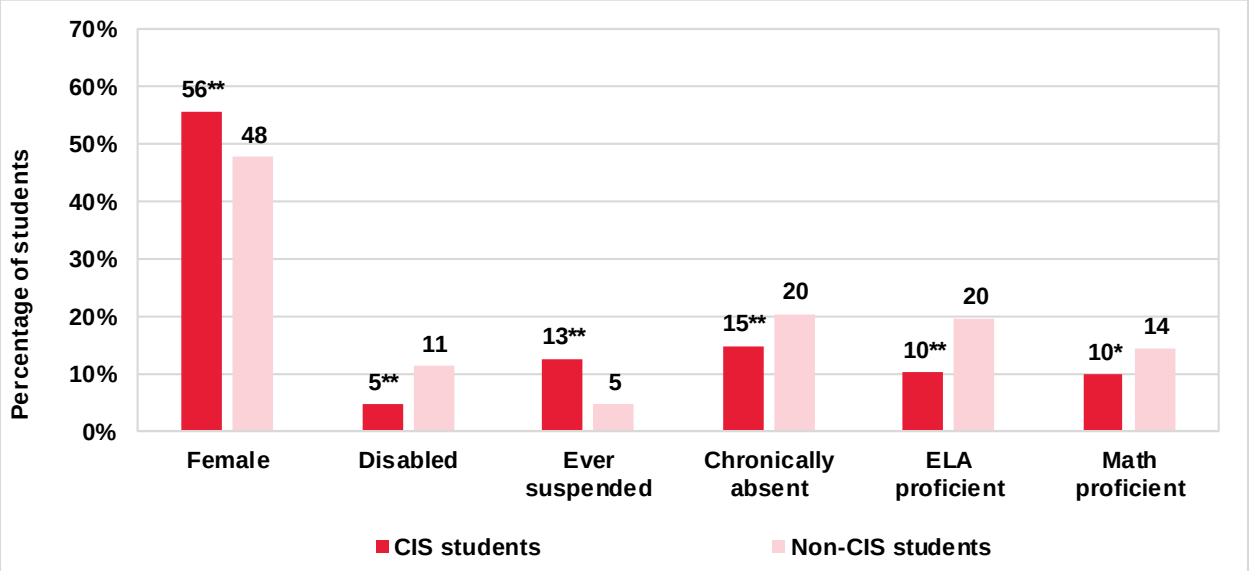
Figure V.1. Number of students on CIS site coordinators' caseloads, by grade level



Source: APS administrative data.

APS = Atlanta Public Schools; CIS = Communities in Schools.

Figure V.2. Baseline characteristics of students on the CIS caseload compared with students not on the CIS caseload in targeted schools



Source: APS administrative data.

Notes: "Ever suspended" shows the percentage of students who were suspended at least one time in the 2016–2017 school year. "Chronically absent" shows the percentage of students who missed 10 percent or more of days enrolled in the 2016–2017 school year. "ELA proficient" and "Math proficient" show the percent of students scoring at least proficient on the STAR fall 2017 assessment for the respective subject.

**Difference is statistically significant at the 1 percent level.

*Difference is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; CIS = Communities in Schools; ELA = English language arts.

B. Impact of CIS case management services

KEY IMPACT FINDING: We found no evidence that the CIS case management component of the Turnaround Strategy improved student suspensions, attendance, or academic achievement.

To analyze the impact of CIS case management services on students' likelihood of being suspended, likelihood of being chronically absent, and assessment scores, we matched students on the CIS caseload with similar students who were not (see Chapter II for more information on the methods and data used). Before receiving case management support from CIS, students on CIS site coordinators' caseloads had baseline outcomes similar to those of the matched comparison group (Table D.4 in Appendix D). Table V.1 summarizes the impacts of receiving CIS case management support. The impacts on nonacademic outcomes are shown in percentage points and the impacts on academic outcomes are shown in standardized *z*-score units.

In spring 2018, after receiving CIS case management support for the school year, there were no statistically significant improvements in caseload students' likelihood of being suspended,[19] likelihood of being chronically absent, or academic performance relative to the matched comparison group (Table V.1).

Table V.1. Impacts of CIS case management on academic and nonacademic student outcomes

	Nonacademic outcomes (percentage points)		Academic outcomes (standard deviations)			
	Likelihood of being suspended	Likelihood of being chronically absent	STAR ELA	STAR Math	GA Milestones ELA	GA Milestones Math
Impact of CIS case management services	0.04	-0.02	-0.02	-0.02	-0.03	-0.03
	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)
Number of students	**1,517**	**1,517**	**1,220**	**1,215**	**1,456**	**1,457**

Source: APS administrative data.

Note: This table displays impact estimates in *z*-scores for the spring 2018 STAR assessment taken by students in all grades and the 2018 Georgia Milestones exams taken by students in grades 3–5 (standard deviations), and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the likelihood that a student was ever suspended during the school year after October 2017. "Chronically absent" refers to the likelihood that a student had missed 10 percent or more of days enrolled. Standard errors are displayed in parentheses below each impact estimate. The sample size reflects the total number of CIS students and matched comparison students in each analysis.

**Impact is statistically significant at the 1 percent level.

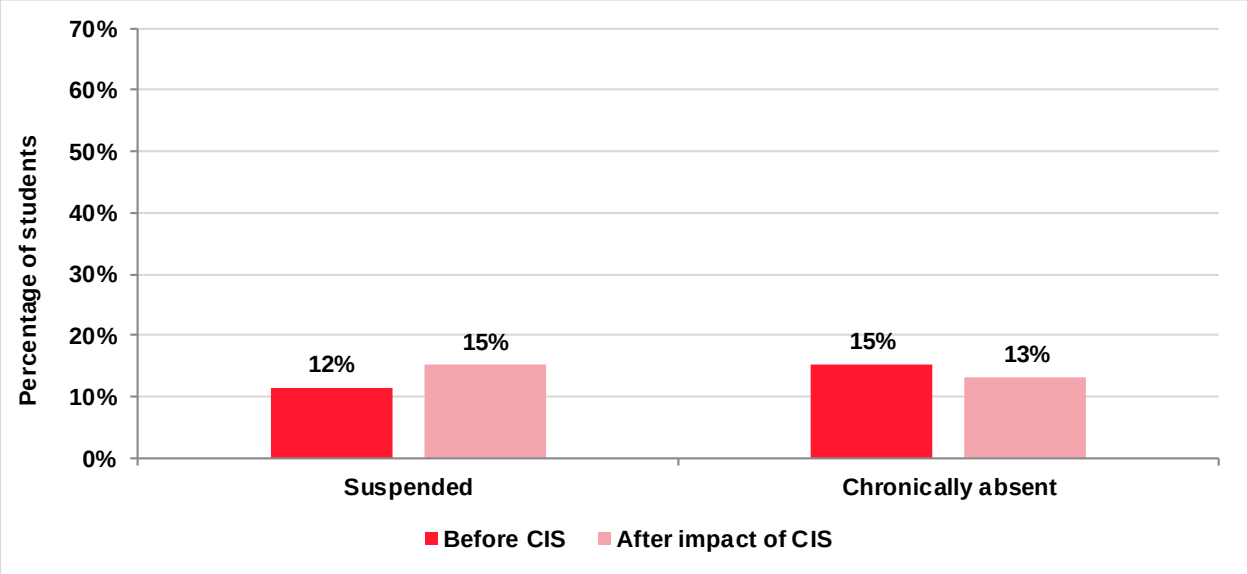
*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; CIS = Communities in Schools; ELA = English language arts.

[19] We examined only suspensions after October 2017, as suspensions in early months may have contributed to students being identified to receive CIS case management support.

Another way to interpret these impact findings is to examine how outcomes changed for students on the CIS caseload as a result of working with a CIS site coordinator. Figure V.3 shows how the percentage of students who were suspended or chronically absent changed from the 2016–2017 to the 2017–2018 school year as a result of receiving support from the CIS site coordinator. There is some indication that the percentage of students suspended increased and the percentage of chronically absent students decreased. However, none of these effects are statistically significant; we cannot say with confidence that CIS case management led to any impacts on students' nonacademic outcomes.

Figure V.3. Impact of receiving CIS case management services on nonacademic outcomes



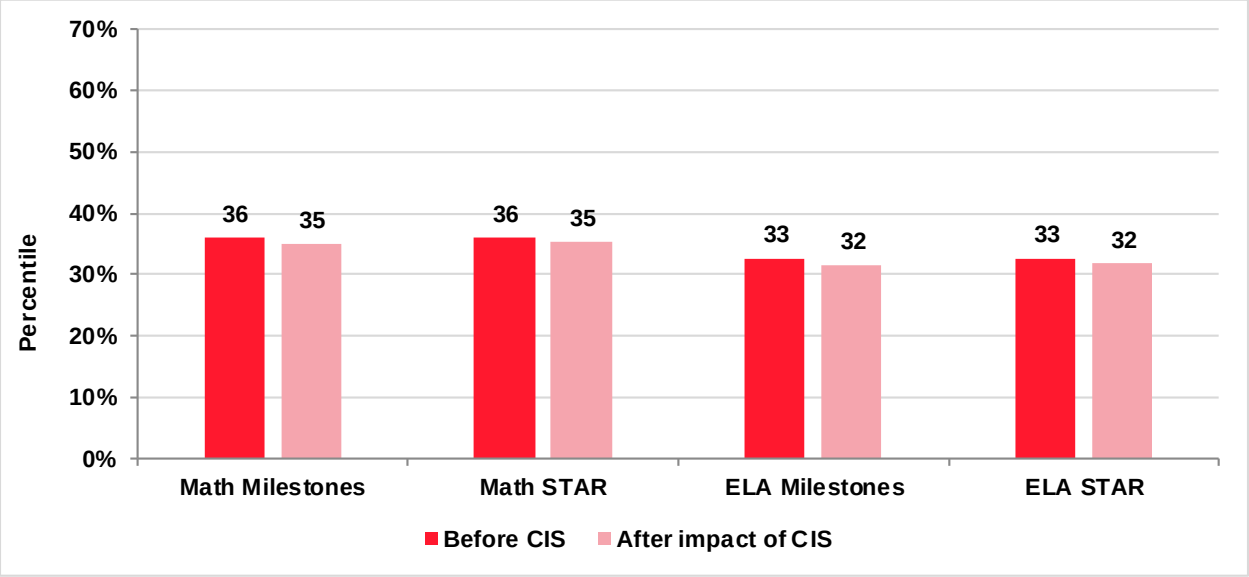
Source: APS administrative data.

Notes: Baseline outcomes are based on the percentage of students suspended and those missing at least 10 percent of days enrolled during the 2016–2017 school year. We calculated the percentages after the impact of CIS by adding the impact estimates to CIS students' baseline percentages. None of the impacts are statistically significant.

APS = Atlanta Public Schools; CIS = Communities in Schools.

Figure V.4 shows how the average test performance of students on the CIS caseload changed from the beginning to the end of the school year as a result of receiving support from the CIS site coordinator, in terms of percentile ranks. The bars indicate a small decline in academic performance in both subjects and assessments, but these impacts are not statistically significant.

Figure V.4. Impact of receiving CIS case management services on academic outcomes



Source: APS administrative data.

Notes: Baseline performance for the percentile ranks before CIS uses the fall 2017 STAR scores for STAR exams and 2017 Georgia Milestones scores for Milestones exams. We calculated the percentile ranks after the impact of CIS by adding the relevant *z*-score impact estimates to CIS students' average baseline *z*-score. We assumed that the percentile rank of the average comparison student did not change over time. None of the impacts are statistically significant.

APS = Atlanta Public Schools; CIS = Communities in Schools; ELA = English language arts.

To better understand the effectiveness of CIS case management, we analyzed whether the impact of CIS case management services differed in schools in which site coordinators had fewer students on their caseloads (and thus could devote more time to each student). We found no evidence that the services' effectiveness differed in schools with smaller caseload sizes (Table D.5 in Appendix D). We also examined whether CIS case management services impacted outcomes for students who were either suspended or chronically absent in the preceding school year (and thus at greater risk in 2017–2018) compared to other CIS students who had not been suspended or chronically absent and thus might be at lower risk. We did not find any evidence that the services' effectiveness differed between these two groups of students (Table D.6 in Appendix D).

C. Implementation of CIS case management services

IMPLEMENTATION CONTEXT

- Staff cited CIS site coordinators' half-time presence and inconsistent schedules as the main challenges to their effectiveness.
- Turnover of CIS staff in some schools may have limited their ability to fully support students.
- CIS case management supports were not highly intensive, and other students in targeted schools may have received more intensive nonacademic supports.

Key challenges to the implementation of the CIS site coordinator role may have limited the impact of case management services on students' outcomes. In addition, other students may have benefited from other potentially helpful nonacademic supports, which could have limited our ability to measure the effectiveness of CIS case management services. These implementation findings may help explain why CIS case management services were not found to be effective.

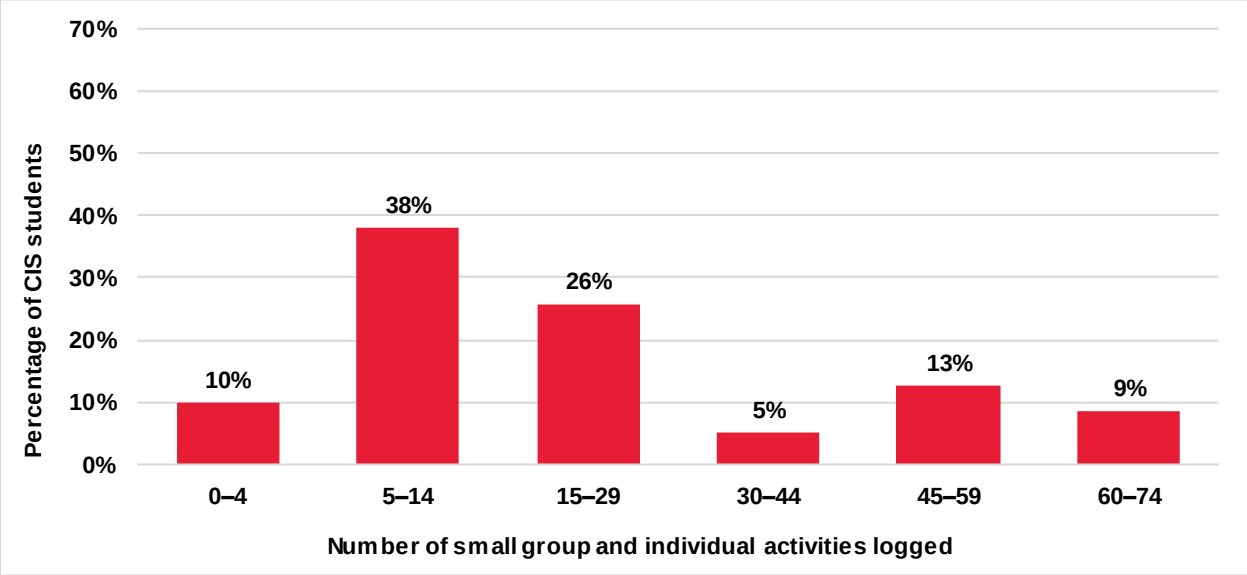
Staff cited CIS site coordinators' half-time presence and inconsistent schedules as the main challenges to their effectiveness. Across all targeted schools, staff perceived CIS site coordinators as only somewhat effective in addressing student behavior. School leaders in several schools stated that a challenge to the success of CIS support was that the coordinators were in schools only two days a week. One respondent explained that the half-time schedule limited opportunities to build consistent and positive relationships with students on their caseload. Staff also noted that because site coordinators split their time between two schools, their schedules were often inconsistent. For example, one respondent reported that the CIS site coordinator did not always adhere to the schedule, which made it difficult to leverage the position because staff did not know when the site coordinator would be at the school.

Turnover of CIS staff in some schools may have limited their ability to fully support students. School leaders at four targeted schools reported that they experienced a turnover of at least one CIS site coordinator assigned to their school over the 2017–2018 school year. The gap between site coordinators at the schools often lasted for several weeks, leaving caseload students without the additional CIS support. Staff at two of the schools explained that, in addition to going weeks without a CIS coordinator, bringing on a new CIS coordinator required building relationships with staff and acquainting the coordinator with the school culture and processes.

CIS case management supports were not highly intensive, and other students in targeted schools may have received more intensive nonacademic supports. Service data that CIS site coordinators logged suggest that CIS case management support was not highly intensive for most students. Almost half of CIS students had fewer than 15 small-group or individual activities logged across the entire school year (Figure V.5). However, other nonacademic supports were also available to students in targeted schools. Each school's Care Team identified students with high nonacademic needs and determined appropriate supports. Some students were placed on the CIS site coordinator's caseload, but others could have worked with another full-time nonacademic staff member, such as a social worker or counselor. One school leader reported ensuring that each of the school's nonacademic support staff members had a caseload of

different students so that the maximum number of students could receive additional support. Data on which students worked with other nonacademic staff or how frequently those staff worked with students were not available, so we were unable to determine how many students received other nonacademic supports or how intensive those services may have been. If students in the comparison group received other nonacademic supports that were effective at improving student outcomes, the effectiveness of CIS case management could be underestimated.

Figure V.5. Percentage of students receiving small-group and individual activities from CIS



Source: APS administrative data.
APS = Atlanta Public Schools; CIS = Communities in Schools.

Our findings are consistent with other recent research on CIS. Parise et al. (2017) examined CIS case management services and found no impact on student attendance or academic performance and a statistically significant increase (that is, an unfavorable effect) in case-managed students' number of suspensions. However, one way in which CIS site coordinators might be helpful to schools is by allowing other full-time support staff to work more closely with other students. Somers and Haider (2017) examined the effect of CIS' whole-school model and found that three years of CIS support led to a positive increase in student attendance in elementary schools, although there were no changes in academic performance. They did not examine the impact of the whole-school model on student suspensions.

In response to feedback from school leaders, APS assigned each targeted school a full-time CIS site coordinator for the 2018–2019 school year. The district may also benefit from collaborating with CIS to determine how to reduce the turnover of site coordinators. The district should consider systematically collecting data on students assigned to work with other nonacademic support staff to assess how intensive and effective these other supports might be. This information could position the district to guide schools on which types of students are better suited to receive support from CIS site coordinators and could help the district decide whether to continue investing in CIS versus other nonacademic supports in the future.

VI. OVERALL FINDINGS FROM PARTNERSHIP SCHOOLS

As part of the Turnaround Strategy in the 2017–2018 school year, partnership organizations managed four APS schools. Purpose Built Schools (PBS) continued its partnership with Thomasville Heights Elementary for a second year and began operating two additional schools—Slater Elementary and Price Middle. Kindezi began operating one school—Gideons Elementary. This chapter briefly describes the model implemented by each partnership organization, analyzes the impact of each on key academic and nonacademic outcomes, and provides implementation findings that give additional context for interpreting the results of the impact analyses.

KEY IMPACT FINDING: Turnaround partnership schools are producing improvements in math performance. However, other effects were mixed, varying by outcome and by partner organization.

A. Partnership with Kindezi

Kindezi operates its schools in accordance with four goals: (1) all students learning, (2) academic ownership, (3) culture and social-emotional learning, and (4) community connectedness. According to the organization, these goals are supported by six pillars: (1) *family-sized classes* that foster opportunities for differentiation and authentic, deep relationships; (2) *excellent teaching* achieved by highly selective hiring, high quality professional development, and career pathways; (3) *challenge and support*, characterized by rigorous expectations accompanied by caring, individualized support; (4) *community and relationships* that have time to build community and connectedness, (5) *racial and socioeconomic diversity* through which all students learn from each other and thrive; and (6) *holistic data-driven*, whereby academic and nonacademic data drive decision making.[20]

In accordance with the six pillars, at Gideons Elementary, Kindezi implemented a small teacher-to-student ratio (one to eight), weekly teacher coaching cycles, after-school programming, and enrichment opportunities for students and staff. The school has a leadership team that includes a principal and multiple assistant principals who also operate as instructional coaches for teachers. Kindezi uses data to monitor the school's progress toward the four goals. For example, school leaders reported that they used student engagement surveys to capture students' motivation to learn and parent engagement surveys to capture feedback from the community. In the next section, we present results from the impact analysis of the Kindezi partnership at Gideons Elementary.

[20] Information on Kindezi's goals and supports can be found at www.kindezi.org.

Impact of Kindezi partnership

KEY IMPACT FINDING: The Kindezi partnership at Gideons Elementary had a positive impact on math and ELA scores, but negatively affected science and social studies scores, and may have increased student suspensions after one year.

We analyzed the impact of the Kindezi partnership at Gideons Elementary using a difference-in-difference design. This design accounts for differences between Gideons and the group of comparison schools, assuming that those differences were consistent over time. In Figure VI.1, we graph standardized average student performance on the state exams[21] in ELA, math, science, and social studies, as well as suspension and chronic absenteeism rates for Gideons Elementary and comparison schools. Because the partnership with Kindezi began in the 2017–2018 school year, Gideons Elementary has had only one year of intervention. It is also important to note that this analysis is based on one school, which makes it impossible to disentangle the effects of Kindezi from any effects specific to Gideons Elementary.

If the Kindezi partnership was effective, we would expect the differences in academic outcomes between Gideons Elementary and comparison schools to shrink after the partnership began. The graphs in Figure VI.1 suggest that this occurred for ELA and math scores, which increased after the partnership. However, science and social studies scores seemed to decline more at Gideons than in the comparison schools. To assess whether the changing trends at Gideons Elementary relative to the comparison schools shown in Figure VI.1 represent real, statistically significant impacts, we conducted a difference-in-difference regression analysis (described earlier and in detail in Chapter II). Table VI.1 presents the impacts of the Kindezi partnership at Gideons Elementary after one year, where zero represents no impact (that is, students at Gideons did as well as would have been expected without Kindezi operating the school).

After one year, the Kindezi partnership led to a statistically significant improvement in students' math and ELA achievement on the Georgia Milestones. However, it also led to a statistically significant decline in Georgia Milestones science and social studies scores. The effects of the Kindezi partnership on academic achievement were roughly equivalent to the following[22]:

[21] Because scales on state tests changed over time, we converted all scaled scores to normalized *z*-scores that show each student's position in the districtwide distribution. Zero represents the districtwide average score; positive scores are above the district average and negative scores are below it. To make the tested grades consistent over time, we analyzed science and social studies performance in grades 5 and 8 only.

[22] This conversion is based on an analysis of annual learning growth on nationally normed exams (Bloom et al. 2008). To convert impacts into months of learning, we divided the impact estimate by the average of the typical annual growth for students in grades 3 through 5 and assumed a nine-month school year. The accuracy of this conversion depends on the extent to which the learning growth on the Milestones exam is similar to the exams analyzed in Bloom et al. (2008). According to this analysis, annual student growth in ELA is 0.60 standard deviations for grade 3, 0.36 for grade 4, and 0.40 for grade 5. Annual student growth in math is 0.89 standard deviations for grade 3, 0.52 for grade 4, and 0.56 for grade 5. Annual student growth in science is 0.40 standard deviations for grade 5 and in social studies is 0.35 for grade 5.

- Students in grades 3 to 5 receiving three additional months of ELA instruction
- Students in grades 3 to 5 receiving five additional months of math instruction
- Students in 5th grade losing seven months of science instruction
- Students in 5th grade losing nine months of social studies instruction

The Kindezi partnership also may have increased the likelihood of a student being suspended by 12 percentage points. This finding is statistically significant and large; however, APS staff and Kindezi staff reported that suspensions at Gideons might have been undercounted before Kindezi's management of the school. If so, our estimate of the unfavorable effect on suspensions would be higher than Kindezi's actual effect. We find no evidence that the Kindezi partnership impacted students' likelihood of being chronically absent.

Table VI.1. Impacts of the Kindezi partnership after one year

	Academic outcomes (standard deviations)				Nonacademic outcomes (percentage points)	
	ELA	Math	Science	Social studies	Suspended	Chronically absent
Impact of Kindezi in Year 1 (Gideons Elementary only)	0.08** (0.04)	0.38** (0.03)	-0.33** (0.07)	-0.41** (0.05)	0.12** (0.01)	-0.01 (0.01)
Number of students	**14,654**	**14,644**	**7,935**	**7,981**	**24,795**	**24,795**
Number of schools	**16**	**16**	**16**	**16**	**16**	**16**

Source: APS administrative data.

Note: This table displays impact estimates for the 2018 Georgia Milestones exams in *z*-scores (standard deviations) and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the likelihood that a student was ever suspended during the school year. "Chronically absent" refers to the likelihood that a student was missing 10 percent or more of days enrolled. Standard errors are displayed in parentheses below each impact estimate. The sample size reflects the total number of students in each analysis across years.

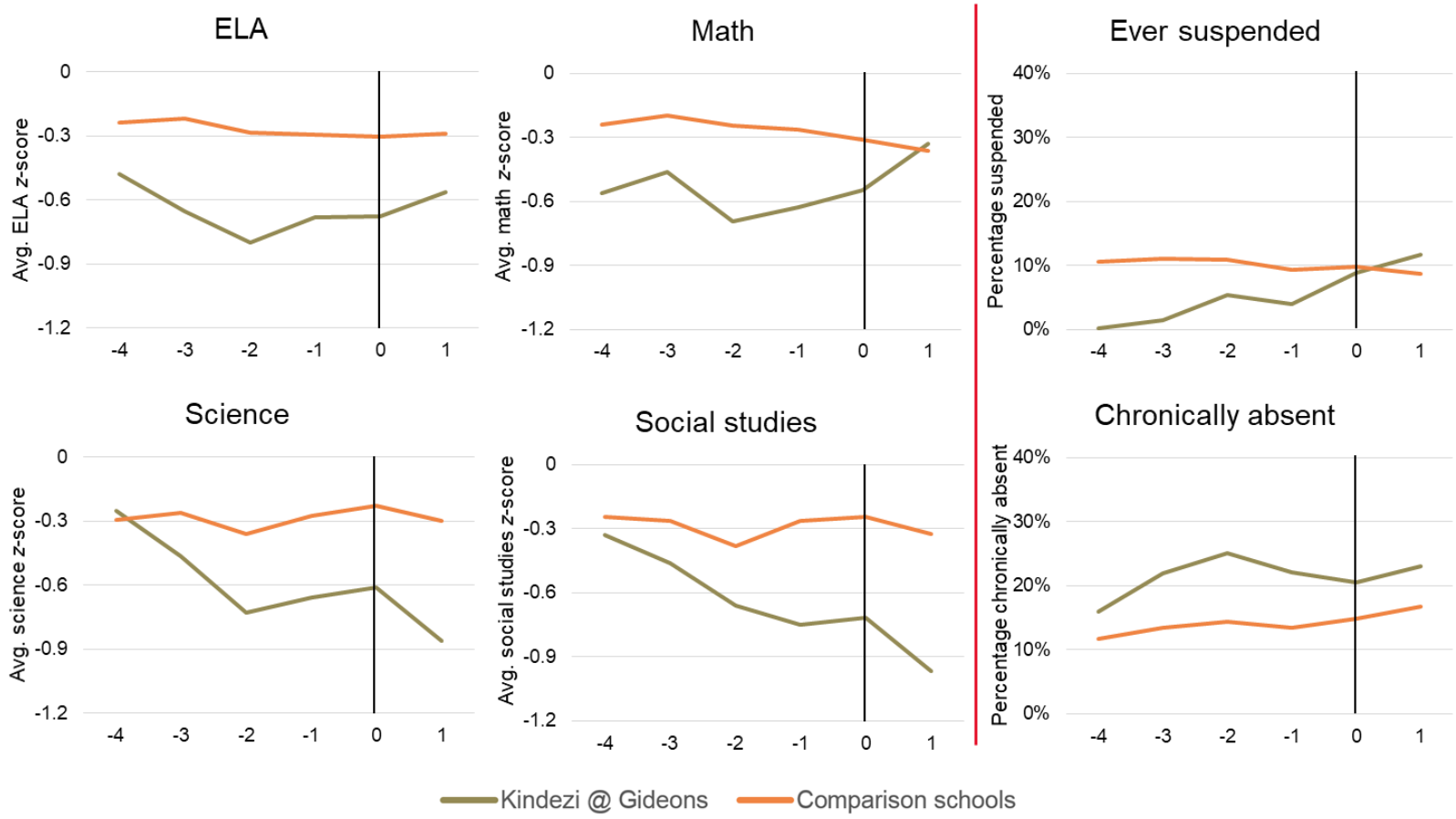
**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; ELA = English language arts.

The impacts on test scores can also be interpreted in terms of how the average student's performance changed after the Kindezi partnership was introduced at Gideons Elementary (Figure VI.2). Academic performance is expressed in terms of the percentile rank of Gideons students on assessments relative to the rest of the district. For example, a student in the 25th percentile performed as well as or better than 25 percent of students in the district. The gray bars on the left indicate actual average performance the year before the Kindezi partnership began. As seen in Table VI.1, there was a positive, statistically significant impact on ELA and math scores and a negative, statistically significant impact on science and social studies scores. These impacts are equivalent to average ELA performance increasing from the 25th to the 28th percentile and the average math performance increasing from the 29th to the 43rd percentile. As

Figure VI.1. Changes in outcomes at Gideons Elementary (operated by Kindezi) and comparison schools over time



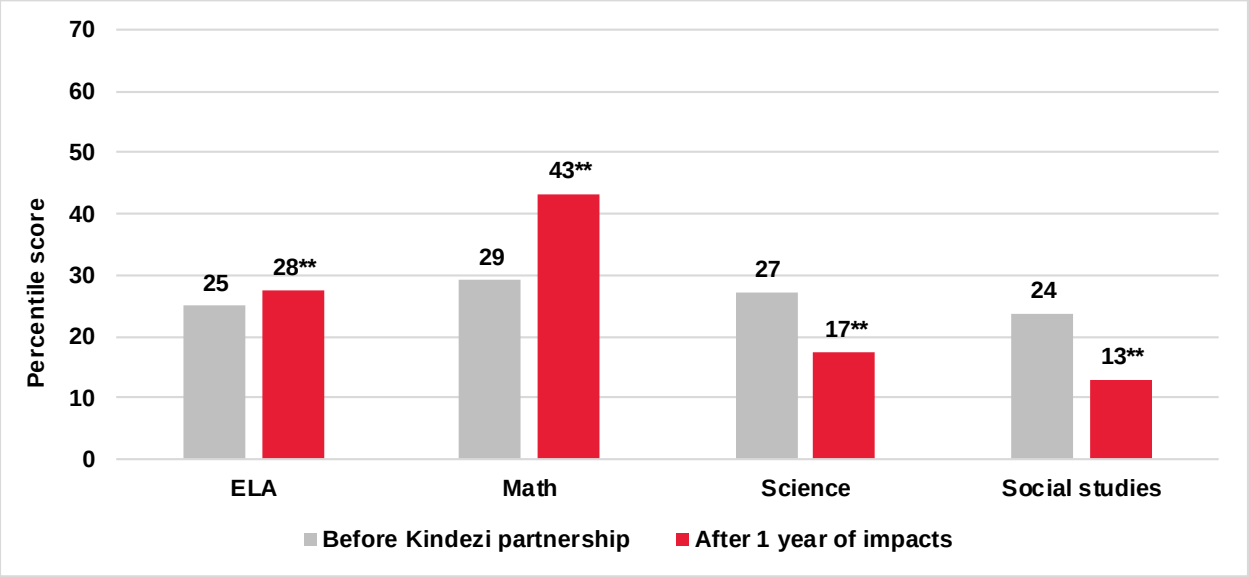
Source: APS administrative data.

Note: These figures display changes in outcome trends for the Georgia Milestones exams in *z*-scores (standard deviations) and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the percentage of students who were ever suspended during the school year. "Chronically absent" refers to the percentage of students missing 10 percent or more of days enrolled. The Kindezi at Gideons trend line shows changes in outcomes across 1 school, and the comparison schools line across 15 schools. Trends tend to be more stable across more schools, as shown in the figure.

APS = Atlanta Public Schools; ELA = English language arts.

a result of the partnership, average performance in science and social studies decreased from the 27th to the 17th percentile and from the 24th to the 13th percentile, respectively.

Figure VI.2. Impacts of the Kindezi partnership on academic outcomes one year after the partnership at Gideons began



Source: APS administrative data.

Notes: We calculated the percentile ranks after one year of the Kindezi partnership by adding the *z*-score impact estimates (Table VI.1) to the average *z*-scores of Gideons Elementary students on the Georgia Milestones exams in the 2016–2017 school year. We assumed that the percentile rank of the average comparison student did not change between years.

**Impact is statistically significant at the 1 percent level.

APS = Atlanta Public Schools; ELA = English language arts.

Implementation of Kindezi partnership

IMPLEMENTATION CONTEXT

- Staff attributed successes in academic achievement to a low student-teacher ratio and supports for teachers.
- Gideons staff modified the school schedule to add time for remediation of foundational reading skills.
- Staff noted improvements in students' behavior but recognized that reported suspensions might increase at the school.
- Despite perceived improvements in behavior, staff recognized the need for additional nonacademic and trauma-informed supports.

We conducted interviews and site visits with Kindezi staff to understand their perceptions of how the partnership was implemented at Gideons Elementary. Included below are key findings that emerged from those conversations, which provide helpful context for interpreting the Kindezi partnership's impact on student outcomes.

Staff attributed successes in academic achievement to a low student-teacher ratio and supports for teachers. Teachers at Gideons felt that academic supports were effective and recognized that the low student-teacher ratio fostered positive relationships with students and opportunities to work individually with those who needed additional support. School leadership reported that they made class size a priority in the budget so that every classroom had two teachers who co-taught approximately 16 students. Each teacher focused on a group of 8 students during instruction. Kindezi leadership also provided Gideons teachers with supports, including lesson plans and intensive professional learning opportunities. Kindezi academic officers developed flexible lesson plans that teachers could use or adapt and that aligned with the scope and sequence of Georgia's state standards. With respect to professional learning opportunities, an assistant principal, who served as an instructional coach, led Kindezi teachers through weekly coaching cycles and planning sessions on making data-driven decisions about which lessons to use.

Gideons staff modified the school schedule to add time for remediation of foundational reading skills. Teachers described significant challenges in implementing the Kindezi curricula because of students' low academic performance (sometimes as low as two to three years behind their grade level) and particular difficulty with foundational reading skills. For this reason, school leadership added 30 minutes to the core academic schedule, dedicated to remediating foundational reading skills. The increased focus on foundational skills may help explain the increase in ELA achievement; however, this focus may have come at the expense of teaching science and social studies content. When describing the daily schedule, teachers reported that science and social studies instruction occurred during the final 45-minute class of the school day, which is also used regularly to deliver social-emotional learning instruction. In contrast, students received approximately two hours of math or ELA instructional content every day of the week in addition to the reading remediation block. It is important to note that only students in grade 5 take the Georgia Milestones exams in science and social studies.[23]

Staff noted improvements in students' behavior but recognized that reported suspensions might increase at the school. School leaders noted that they felt behavioral issues improved compared with their observations at the school before the Kindezi partnership. They attributed these successes to the range of nonacademic supports offered at the school, which included an on-site social worker, a counselor, and four behavioral aides. Gideons staff also reported success in progressing toward using in-school rather than out-of-school suspensions to keep students accountable and minimize loss of instructional time, but noted that there were times when out-of-school suspensions were necessary. For example, staff explained that several staff who managed discipline went on approved leave for an extended period of time during the year and out-of-school suspensions became the primary discipline strategy during that time.

In addition, during interviews staff stated that they began reporting suspensions more accurately compared with previous years when the school was not under Kindezi management, and anticipated an increase in the reported suspension rate from previous years. As a result, the

[23] We examined the ELA and math impacts of the Kindezi partnership in grade 5 only and did not find large declines in those subjects that mirrored the pattern in science and social studies. This finding suggests that grade-specific factors do not wholly explain the findings. (Results not shown.)

impact on suspensions could at least partly reflect underreporting of suspensions before the Kindezi partnership.

Despite perceived improvements in behavior, staff recognized the need for additional nonacademic and trauma-informed supports. Although staff acknowledged behavioral improvements, several teachers reported persistent behavioral issues from students, including fighting and classroom disruptions that impeded student learning. To address classroom management issues, teachers said they needed additional supports, such as additional behavioral aides or stronger expectations about behavior and stronger consequences for students falling short of expectations. Kindezi leadership also acknowledged that students at Gideons had a much higher need for nonacademic services compared with students at the other Kindezi charter schools in APS. They described ongoing efforts to form partnerships with organizations that have the capacity to address those needs, such as those offering wraparound services and mental health support. Kindezi leaders plan to provide individualized professional development to school leaders and teachers on trauma-informed care for students. These efforts may enable the school to more effectively support students' needs.

Taking the implementation findings into account, the impacts on academic scores are likely explained by the Kindezi model's heightened focus on ELA and math instruction, which may have come at the expense of science and social studies instruction, whereas the apparent increase in suspensions could be a result of changes in reporting practices. Ongoing efforts to address the higher level of nonacademic needs among Gideons students compared to those in other Kindezi schools may further drive improvements.

B. Partnerships with Purpose Built Schools

PBS schools follow the Drew Instructional Model, which focuses on high quality and increased instructional time, as well as comprehensive systems of student and family support. According to the organization, PBS schools emphasize literacy and math instruction and supports to promote student learning; rely on high quality teachers; and use a thematic, project-based curriculum focusing on science, technology, engineering, arts, and mathematics (STEAM). PBS schools implement several supports to address students' academic needs, including an extended school day, afterschool programs, enrichment opportunities, and pre-kindergarten in elementary schools. The PBS model also emphasizes early intervention provided through a math and literacy lab where two teachers and a specialist provide supplemental academic support to the lowest-performing students in their subject area through one-on-one or small group instruction.[24]

PBS staff reported that their schools also feature a number of family and nonacademic supports. Each school has at least one family and community outreach coordinator who monitors attendance, conducts outreach to families, and works with referred students and families to address their needs. PBS schools communicate with families through their phone messaging application and they have a parent group in which parent representatives communicate with other parents. They also offer a number of supports to families, such as legal services from the Atlanta

[24] Information on the PBS model can be found at www.purposebuiltschoolsatlanta.org.

Volunteer Lawyers Foundation and wraparound services and training for their staff through CHRIS 180, a nonprofit organization with expertise in trauma-informed care.

PBS has an administrative leadership team that works across its schools in the APS system. In addition, each elementary school has a leadership team that includes a principal, an assistant principal, and instructional coaches. At Price Middle School, the administrative team is composed of a principal, a dean of students, a dean of academics, and an instructional coach. Each school also has several support staff members, resulting in approximately one adult for every five students at the school. In addition to regular classroom teachers, PBS schools have "connections teachers" who oversee band, robotics, STEAM, Spanish, and other specialized class offerings. Additional staff members work as co-teachers or paraprofessionals in regular education and special education classrooms, or provide instructional support in ELA and math labs. Support staff members, such as the Response to Intervention (RTI)/Student Support Team (SST) specialist and gifted teacher, offer supplemental academic supports to students; other staff, including the counselor and family outreach coordinators, provide individual and socioemotional supports to address the nonacademic needs of students and their families.

Impact of PBS partnership

KEY IMPACT FINDING: The PBS partnership had a positive impact on math scores after one and two years. However, the PBS partnership also saw increased student suspensions and had inconsistent effects on chronic absenteeism.

We analyzed the impact of the PBS partnership at Thomasville Heights Elementary, Slater Elementary, and Price Middle using the same strategy we used in analyzing the Kindezi partnership. In Figure VI.3, we graph standardized average student performance on the state exams in ELA, math, science, and social studies, as well as suspension and chronic absenteeism rates, for the three PBS partnership schools and the comparison schools. The outcomes are shown for two years following the start of the partnership at Thomasville Heights, which became a PBS partnership school in the 2016–2017 school year, and one year following the start of the partnership at Slater and Price, which became PBS partnership schools in the 2017–2018 school year. To measure whether the changing trends in PBS schools relative to the comparison schools illustrated in Figure IV.3 represent real, statistically significant impacts, we conducted a difference-in-difference regression analysis (described earlier and in detail in Chapter II).[25]

We estimated the impact of one year of the PBS partnership on each outcome across all three partnership schools.[26] Because PBS began operating Thomasville Heights in the 2016–2017

[25] Although math scores decreased from prior years in the first year of the partnership at Slater Elementary and Price Middle and in the second year at Thomasville Heights Elementary, they also decreased in comparison schools during the same period. The results of the regression analysis indicate that there were positive impacts in math in both years despite this downward trend.

[26] Our previous impact report (Hallgren et al. 2017) includes an impact analysis of one year of the PBS partnership at Thomasville Heights alone. That study found that the PBS partnership at that school improved student achievement in math and social studies after one year, but had little effect on ELA scores, science scores, or student absences. That study did not examine student suspensions.

school year, we also estimated the impact of the PBS partnership after two years at that school. It is important to note that the results of that analysis are based on one school, which makes it impossible to disentangle the effects of PBS after two years from any effects specific to Thomasville Heights. Table VI.2 presents the measured impacts of the PBS partnership in the first and second years.

The results of the impact analysis indicate mixed success for the PBS partnership. Across all three PBS partnership schools, one year of PBS supports led to the following changes in students' outcomes:

- Students' math performance on the Georgia Milestones exam improved by an amount roughly equal to almost two additional months of math instruction.
- Students' likelihood of being chronically absent decreased by three percentage points.
- Students' likelihood of being suspended may have increased by eight percentage points.

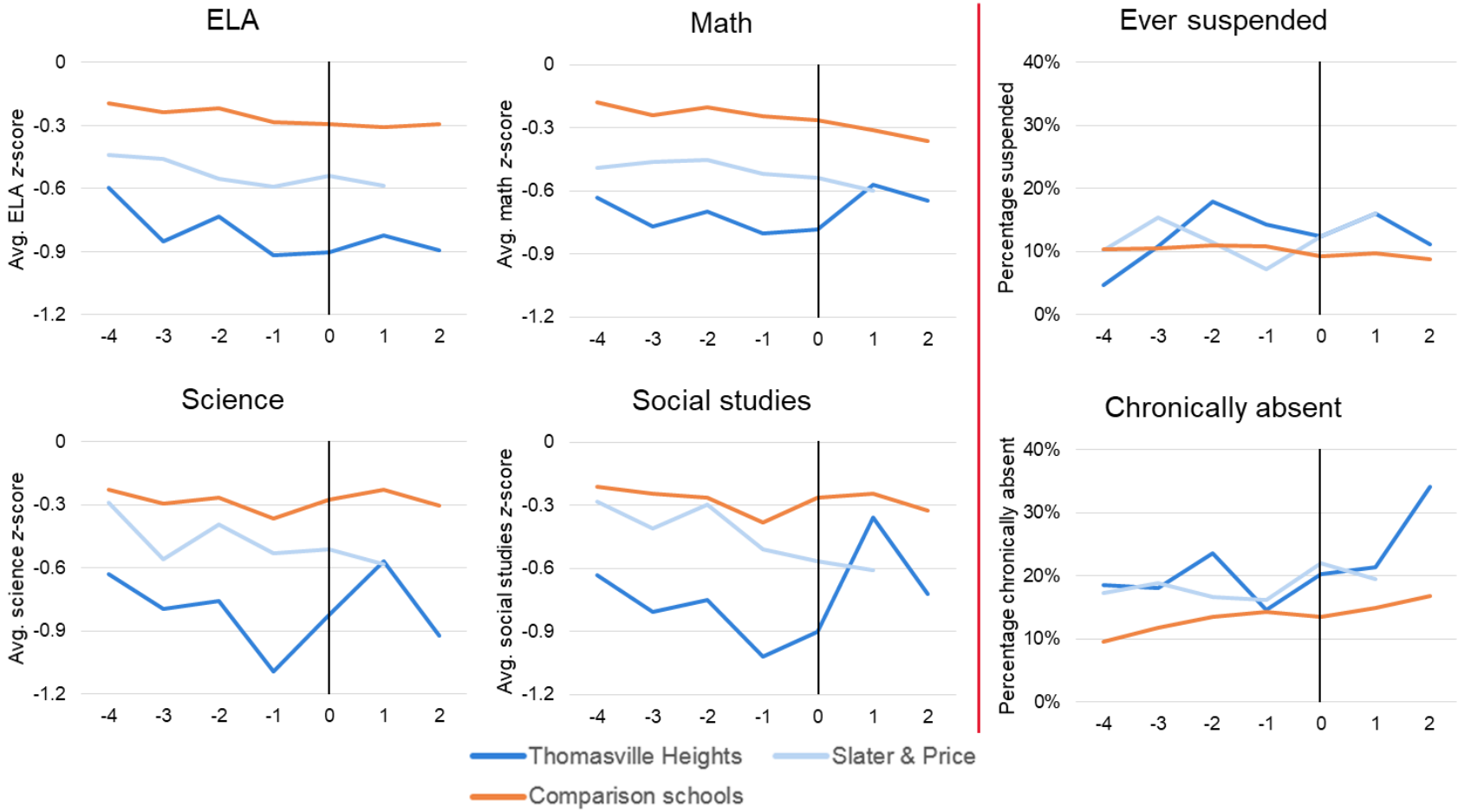
At Thomasville Heights, the second year of the partnership led to the following changes in students' outcomes:

- Students' math performance on the Georgia Milestones exam improved by an amount roughly equal to four additional months of math instruction.
- Students' likelihood of being chronically absent increased by eight percentage points.
- Students' likelihood of being suspended may have increased by three percentage points.

Each of these results is statistically significant. Estimated impacts on suspensions should be interpreted with caution, however, because APS staff indicated that suspensions may not have been recorded accurately before PBS management. If suspensions were previously underreported, our estimate of PBS' unfavorable effect on suspensions would be higher than the actual effect. The partnership did not lead to a statistically significant impact on students' ELA, science, or social studies scores in either year.

The impacts on test scores can also be interpreted in terms of changes in students' percentile scores after the PBS partnership was introduced at the schools. Figure VI.4 shows how the average performance of students at Thomasville Heights, Slater, and Price changed following one year of impacts, as well as how the average performance of students at Thomasville Heights changed following two years of impacts. The gray bars on the left indicate actual average performance the year before the PBS partnership began at the schools in terms of the percentile rank of students in the schools relative to the rest of the district. As seen in Table VI.2, there was a positive, statistically significant impact on math scores in both years. These impacts are equivalent to average math performance increasing from the 28th to the 32nd percentile after one year of supports and from the 28th to the 42nd percentile after two years of supports. There is some indication that science and social studies performance decreased, but these impacts were not statistically significant.

Figure VI.3. Changes in outcomes at PBS partnership schools and comparison schools over time



Source: APS administrative data.

Note: These figures display changes in outcome trends for the Georgia Milestones exams in *z*-scores (standard deviations) and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the percentage of students who were ever suspended during the school year. "Chronically absent" refers to the percentage of students missing 10 percent or more of days enrolled. The Thomasville Heights trend line shows changes in outcomes across 1 school; the Slater & Price line across 2 schools; and the comparison schools line across 15 schools. Trends tend to be more stable across more schools, as shown in the figure.

APS = Atlanta Public Schools; ELA = English language arts; PBS = Purpose Built Schools.

Table VI.2. Impacts of PBS partnerships after one and two years

	Academic outcomes (standard deviations)				Nonacademic outcomes (percentage points)	
	ELA	Math	Science	Social studies	Likelihood of being suspended	Likelihood of being chronically absent
Impact of PBS in Year 1	-0.01 (0.03)	0.12* (0.04)	-0.08 (0.15)	-0.14 (0.19)	0.08** (0.01)	-0.03** (0.01)
Impact of PBS in Year 2 (Thomasville Heights only)	-0.06 (0.03)	0.26** (0.03)	-0.13 (0.07)	0.02 (0.06)	0.03** (0.01)	0.08** (0.01)
Number of students	**15,913**	**15,908**	**8,781**	**8,828**	**26,692**	**26,692**
Number of schools	**18**	**18**	**18**	**18**	**18**	**18**

Source: APS administrative data.

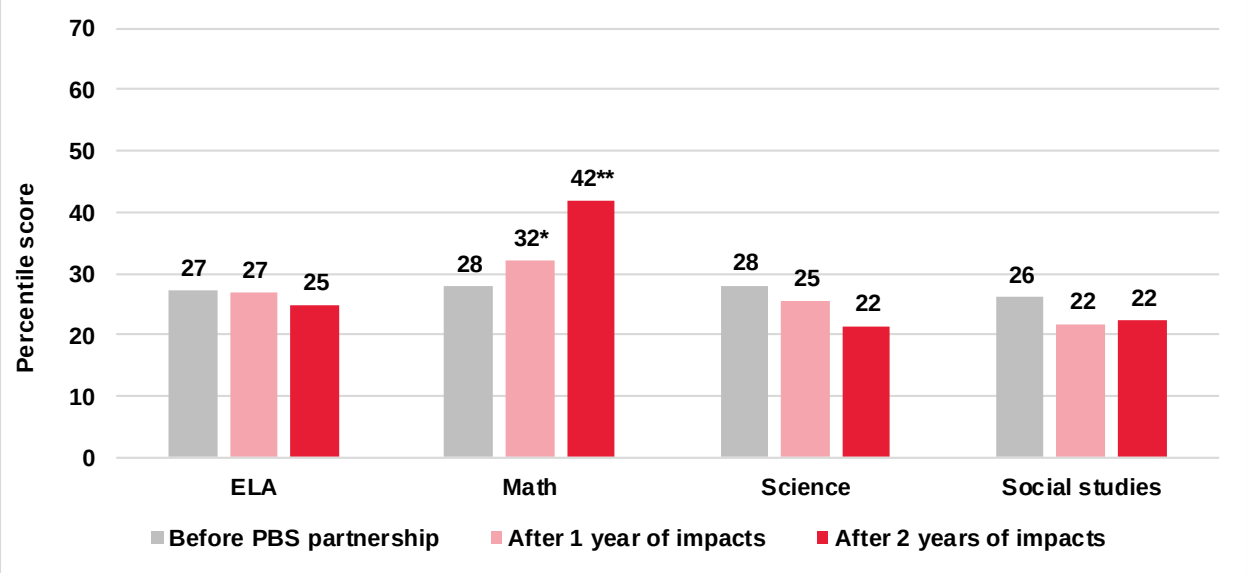
Note: This table displays impact estimates for the 2017 and 2018 Georgia Milestones exams in *z*-scores (standard deviations) and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the likelihood that a student was ever suspended during the school year. "Chronically absent" refers to the likelihood that a student was missing 10 percent or more of days enrolled. Standard errors are displayed in parentheses below each impact estimate. The sample size reflects the total number of students in each analysis across years.

**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; ELA = English Language Arts; PBS = Purpose Built Schools.

Figure VI.4. Impacts of the PBS partnership on academic outcomes one and two years after the partnership began



Source: APS administrative data.

Notes: We calculated the percentile ranks after one and two years of the PBS partnership by adding the *z*-score impact estimates (Table VI.2) to the average *z*-scores of PBS partnership students on the Georgia Milestones exams in the school year before partnerships began. We assumed that the percentile rank of the average comparison student did not change between years.

**Impact is statistically significant at the 1 percent level.

APS = Atlanta Public Schools; ELA = English language arts; PBS = Purpose Built Schools.

Implementation of PBS partnership

IMPLEMENTATION CONTEXT

- Staff attributed success in academic achievement at PBS schools to a generous student-teacher ratio.
- School leaders made substantial modifications to new academic curricula to support students below grade level.
- PBS staff, particularly at Thomasville Heights, noted that mid-year integration of restorative justice practices supported perceived behavior changes.

We conducted interviews and site visits with PBS staff to understand their perceptions of how the three PBS schools implemented the partnership. Included below are key findings that emerged from those conversations, which provide helpful context to interpret the PBS partnership's impact on student outcomes.

Staff attributed success in academic achievement at PBS schools to a generous student-teacher ratio. Staff across PBS partnership schools believed that the partnership was very effective in improving students' academic achievement. Teachers attributed this perceived success in academic achievement to the generous adult-to-student ratio, designed to provide students with access to more staff who offered personalized attention and socioemotional supports. This ratio exists in each classroom in which teachers work alongside paraprofessionals, and in the literacy and math labs in which multiple teachers provide instruction and support.

School leaders made substantial modifications to new academic curricula to support students below grade level. The three PBS schools adopted new academic curricula for the 2017–2018 school year. Slater and Price adopted new curricula across all content areas; Thomasville Heights adopted a new ELA curriculum. According to school leaders, staff needed to heavily modify the new curricula to meet their students' needs, create new lesson plans, and provide additional training to teachers. School leaders also described the need to substantially supplement the curricula. For example, staff at Thomasville Heights used extensive supplemental materials aimed at building students' foundational skills to accompany the school's new ELA curriculum, which assumed students already had these skills. Staff described the transition to new curricula as difficult because of the increased rigor of the material. They explained that before the PBS partnership, students had grown accustomed to a less demanding curriculum. Staff described how, in the face of the new challenging content and more rigorous expectations, students acted out and disrupted the class, would not engage, or refused to continue their classwork. Despite this reaction, staff stated that they continued to push students and implemented a number of instructional strategies, such as breaking down difficult concepts over multiple activities, using academic games, having students rotate through small group instruction, and using manipulatives. Staff described how over time they felt students became "comfortable with growing," appreciated the rigor of the curriculum, and demonstrated positive academic changes.

PBS staff, particularly at Thomasville Heights, noted that mid-year integration of restorative justice practices supported perceived behavior changes. Staff in all three schools experienced significant behavioral issues with students; in response, they adopted restorative justice practices mid-way through the 2017–2018 school year. CHRIS 180 provided training on these practices, which are intended to avoid punitive measures and infuse safety and trust into the school culture. Staff used restorative justice practices to manage disputes or behavioral issues and encouraged students to use them as strategies for expressing frustration or other negative feelings.

Staff at Price and Slater said they noticed positive changes after introducing restorative justice practices, but noted that student behavior posed an ongoing challenge. They attributed behavior issues to the trauma students experienced outside of school and characterized it as part of working in a turnaround school. Staff at Price, in particular, explained that students' familiarity with how the school operated before the PBS partnership created resistance to the behavior expectations imposed by the PBS model, adding friction to the relationship between students and teachers. Staff at Thomasville Heights, on the other hand, described an immediate improvement in student behavior after integrating restorative justice practices in the winter of 2017–2018 (the second year of the partnership at the school). They noted, for example, that students started using restorative justice language with staff and their peers without prompting. Thomasville Heights staff felt that the school culture grew stronger as students began trusting staff, recognized the school as a safe space, and established mutual respect between students and teachers.

The favorable impressions of PBS staff regarding the effects of restorative justice practices may be at odds with our findings on suspensions, which suggest that suspension rates increased under the PBS partnerships—but, as we noted earlier, this unfavorable impact may be overstated if suspensions were underreported before PBS management. Even so, PBS' practices for reporting suspensions should have been consistent throughout the 2017–2018 school year, and we found no evidence that suspensions declined in the second half of the year, after restorative justice practices were implemented.

These implementation findings help interpret the findings from the impact analysis. Changes introduced by PBS, such as more rigorous curricula and intensive supports for low-performing students, may explain the improvements in students' math performance. An analysis of the partnership after the 2018–2019 school year will provide insight into the extent that integration of restorative justice practices at PBS schools reduced suspension referrals and will assess whether student performance in other subjects improved. In addition, PBS began operating an additional partnership school at Carver High School in the 2018–2019 school year.

The final year of the study will include an updated analysis of partnership schools, which will examine the effectiveness of the Kindezi partnership at Gideons Elementary after two years of supports as well as the impact of the PBS partnerships across all four PBS partnership schools.

VII. CONCLUSION

This study of the APS Turnaround Strategy contributes to the growing body of evidence on improving low-performing schools. Consistent with findings from other studies of turnaround efforts, two years of Turnaround Strategy supports in a subset of APS schools resulted in some improvements as well as recognition of challenges faced. This concluding chapter presents a synthesis of findings from the second year of the study, recommendations to assist APS in determining how best to support the Turnaround Strategy going forward, and next steps for the evaluation.

A. Synthesis of findings

The Turnaround Strategy is producing improvements in math performance in both targeted and partnership schools. Targeted supports, the Kindezi partnership, and the PBS partnership all improved students' math scores on the Georgia Milestones. This finding is a promising sign that the Strategy is supporting growth in students' academic performance in ways that would not have been achieved without it. We cannot, however, determine with certainty what has driven these improvements. The analyses performed capture the effect of the comprehensive set of supports that the district provides to targeted schools or that Kindezi or PBS offers to their partnership schools. After two years of the evaluation, we have not found positive impacts from Strategy components aimed at individual students, such as High Impact Tutoring or math and reading specialists, so it is likely that schoolwide changes are driving the positive math impacts.

The limited impacts on other outcomes reflect the challenges of successfully improving schools and are consistent with much of the school turnaround literature. We did not find consistent evidence of positive impacts of schoolwide targeted support or school partnerships on other student outcomes. These findings are consistent with other studies of school turnaround efforts. For example, three years of turnaround efforts in Houston resulted in an increase in math achievement but had no effect on reading achievement or attendance in elementary schools (Fryer 2014). An evaluation of Project LIFT, a five-year turnaround effort in the Charlotte-Mecklenburg School District, did not find positive effects of the supports on elementary students' math or reading scores (Research for Action 2018). In Tennessee, which implemented three different turnaround models to transform low-performing schools, only one of the three models succeeded in improving student outcomes (Zimmer et al. 2017). Also, a national study of turnaround efforts implemented under federal School Improvement Grants found no impacts on student achievement on average (Dragoset et al. 2017). Continued research about the complexity and nuances of improvement efforts will be helpful for developing a deeper understanding of effective ways to improve schools in need of support.

Specialists did not significantly impact the academic performance of students on their rosters, but took on additional responsibilities in the schools that may have provided other value. We did not find strong evidence that math and reading specialists impacted Georgia Milestones or STAR scores for students on their rosters. Even so, staff in targeted schools saw the specialists as highly valuable additions to their schools, and the schools used the specialists in a variety of ways. Specialists frequently worked with students not on their rosters and supported teacher professional learning in an effort to improve instruction in the schools. Because some

students who worked with specialists were not on their rosters, and because specialists may have produced some schoolwide effects, our analysis may have underestimated specialists' true impact. Moreover, there is suggestive evidence that specialists may be effective in improving the achievement of especially low-performing students—the population the specialists were originally expected to serve. A renewed focus on those students might produce better results.

The limited impacts of CIS case management services are consistent with other recent studies, and may be partly explained by implementation challenges. Receiving small group or individual support from a CIS site coordinator did not significantly impact students' likelihood of being chronically absent, likelihood of being suspended, or performance on Georgia Milestones exams. The limited success of this role may be because of implementation challenges, such as most schools only having one CIS site coordinator at their campus two days a week, high turnover among CIS staff, or the low intensity of small group and individual support that most students experienced. However, these results are in line with other recent research on CIS case management services (Parise et al. 2017). It is possible that the addition of a CIS site coordinator in targeted schools allowed other nonacademic support staff, such as clinical psychologists and social workers, to provide more support to other students in the schools than would have been possible if the CIS site coordinator was not present.

B. Recommendations

APS should continue to explore how to support students' growth in subjects other than math. The Kindezi partnership at Gideons Elementary improved students' performance in ELA; however, supports in targeted schools and the PBS partnership have not yet shown improvements in students' ELA achievement. Kindezi strongly emphasized and resourced remediating students' foundational reading skills, and added substantial time for reading instruction in the school's daily schedule. Although both targeted and PBS schools have additional staff to support ELA instruction, an even greater focus on remediating foundational reading skills may be necessary. Research shows that it is typically more challenging to improve ELA performance than math performance (see, for example, Fryer 2014).

There are also indications that schools should pay attention to science and social studies instruction. The Kindezi partnership significantly worsened students' performance on the science and social studies Georgia Milestones even while it improved outcomes in math and ELA. Science and social studies performance also declined in targeted and PBS schools, although those impacts were not statistically significant. This finding suggests that there may be a trade-off when improving test scores across different subjects. Specifically, as schools devote additional time to remediating foundational skills, there may be less time remaining for science and social studies instruction.

The district could capture richer program data to better understand which supports are most effective, and for whom. Specifically, capturing the frequency and duration that math and reading specialists meet with students on their rosters and regularly updating those rosters would position the district to better understand the specialists' effectiveness. Similarly, tracking which students work with nonacademic support staff could enable the district to assess the effectiveness of the different types of nonacademic supports available in addition to CIS and identify which students could benefit most from each type of support.

The district should monitor the use of suspensions, particularly in partnership schools. The reported number of student suspensions increased in both the PBS and Kindezi partnerships. This finding should be interpreted cautiously, because APS and partnership staff reported that suspensions may have been underreported before the partnerships began (in which case the apparent increase may not be real). Staff from each partnership reported improvements in student behavior but recognized that behavior continued to be a challenge in the schools. In addition, although we do not find that targeted supports led to increased suspensions, staff in those schools also reported behavior issues.

Despite efforts to address behavior challenges, suspension rates in partnership and targeted schools remained relatively consistent over the course of the 2017–2018 school year. Each partnership school reported efforts to address behavior: Kindezi used in-school discipline techniques rather than suspensions whenever possible, and PBS introduced restorative justice practices at mid-year. These strategies could lead to improved results in future years, but we find no evidence that they had reduced suspension rates by the end of the 2017–2018 school year. APS and the partnership organizations should closely monitor suspension rates to assess whether an improving trend becomes evident during the 2018–2019 school year.

Extending the supports in the Turnaround Strategy may be critical as schools begin to improve. Targeted and partnership schools reported that they saw an improvement in students' academics and nonacademics, yet they also pointed to the need for additional and continued improvements in both areas. For example, school staff described how students made gains in academic achievement but noted that many students still did not perform at grade level. School staff also noted that students showed growth in their social and emotional skills but continued to behave in ways that disrupted student learning during class time. APS may want to consider ways of offering academic and nonacademic supports to even more students at low-performing schools. In addition, targeted supports led to improvements in math achievement after two years of supports, but there were no improvements after only one year. This finding is consistent with other turnaround literature that highlights how turning around low-performing schools often takes several years (Chin et al. 2018). As schools begin to show improvements, APS might consider how to scaffold or extend the Turnaround Strategy supports so schools can effect lasting change.

C. Next steps for the evaluation

The final year of the evaluation will include an implementation study of the third year of Turnaround Strategy supports as well as additional impact analyses. The evaluation team will collaborate with APS to determine how to focus our analyses in the study's third and final year. For example, we may examine critical factors of turnaround success, such as school culture, leadership abilities, or staffing mobility and quality (see Center on School Turnaround 2017). The information gained from the study should help APS and other school districts learn more about ways to effectively support low-performing schools.

REFERENCES

Benjamini, Y., and Y. Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society Series B Methodological*, vol. 57, no. 1, 1995, pp. 289–300.

Bloom, H.S., C.J. Hill, A.R. Black, and M.W. Lipsey. "Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions." *Journal of Research on Educational Effectiveness*, vol. 1, no. 4, 2008, pp. 289–328.

The Center on School Turnaround. "Four Domains for Rapid School Improvement: A Systems Framework." 2017. San Francisco, CA: WestEd. Available at http://centeronschoolturnaround.org/wp-content/uploads/2017/02/CST_Four-Domains-Framework-Final.pdf.

Chin, Mark, Thomas J. Kane, Whitney Kozakowski, Beth E. Schueler, and Douglas O. Staiger. "School District Reform in Newark: Within- and Between-School Changes in Achievement Growth." *ILR Review*, 2018. doi: 10.1177/0019793918774318.

Dougherty, Shaun M. and Jennie M. Weiner. "The Rhode to Turnaround: The Impact of Waivers to No Child Left Behind on School Performance." *Educational Policy*, 2017. doi: 10.1177/0895904817719520.

Dragoset, L., J. Thomas, M. Herrmann, J. Deke, S. James-Burdumy, C. Graczewski, A. Boyle, R. Upton, C. Tanenbaum, and J. Griffin. "School Improvement Grants: Implementation and Effectiveness." NCEE 2017 4013. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2017.

Fryer Jr., Roland G. "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments." *Quarterly Journal of Economics*, vol. 129, no. 3, 2014, pp. 1355–1407.

Gill, B., J. Furgeson, H. Chiang, B. Teh, J. Haimson, and N. Verbitsky Savitz. "Replicating Experimental Impact Estimates with Nonexperimental Methods in the Context of Control-Group Noncompliance." *Statistics and Public Policy*, vol. 3, no. 1, 2015.

Gill, Brian, Ron Zimmer, Jolley Christman, and Suzanne Blanc. "Student Achievement in Privately Managed and District-Managed Schools in Philadelphia Since the State Takeover." Santa Monica, CA: RAND Corporation, 2007. Available at https://www.rand.org/pubs/research_briefs/RB9239.html.

Hallgren, Kristin, Jane Choi, Annie Li, Lindsay Ochoa, and Kevin Kelly, Lindsay Ochoa. "Implementation Findings from Year 2 of the Atlanta Public Schools' Turnaround Strategy Evaluation." Report submitted to the Atlanta Public Schools. Princeton, NJ: Mathematica Policy Research, July 2018.

Hallgren, Kristin, Naihobe Gonzalez, Jane Choi, Kevin Kelly, Lindsay Ochoa, and Ann Li. "The Atlanta Public Schools Turnaround Strategy After One Year: High Impact Tutoring and the Purpose Built Schools Partnership." Report submitted to the Atlanta Public Schools. Princeton, NJ: Mathematica Policy Research, October 2017.

Heissel, Jennifer A., and Helen F. Ladd. "School turnaround in North Carolina: A Regression Discontinuity Analysis." *Economics of Education Review*, vol. 62, 2018, pp. 302–320. Available at https://www.sciencedirect.com/science/article/pii/S0272775716303156. Accessed January 6, 2019.

The Kindezi Schools. 2019. Available at www.kindezi.org. Accessed January 4, 2019.

Parise, Leigh, William Corrin, Kelly Granito, Zeest Haider, Marie-Andrée Somers, and Oscar Cerna. "Two Years of Case Management: Final Findings from the Communities in Schools Random Assignment Evaluation." MDRC, 2017.

Purpose Built Schools Atlanta. 2019. Available at http://purposebuiltschoolsatlanta.org/. Accessed January 4, 2019.

Research for Action. "Project LIFT: Year Five Student Outcomes Memo." 2018. Available at http://www.projectliftcharlotte.org/wp-content/uploads/2018/02/LIFT-Year-5-student-outcomes-memo-and-executive-summary.pdf. Accessed January 6, 2019.

Somers, Marie-Andrée, and Zeest Haider. "Using Integrated Students Supports to Keep Kids in School: A Quasi-Experimental Evaluation of Communities in Schools." MDRC, 2017.

Tuttle, C., B. Gill, P. Gleason, V. Knechtel, I. Nichols-Barrer, and A. Resch. "KIPP Middle Schools: Impacts on Achievement and Other Outcomes." A Report of the National Evaluation of KIPP Middle Schools. Washington, DC: Mathematica Policy Research, 2013.

Zimmer, R., G.T. Henry, and A. Kho. "The Effects of School Turnaround in Tennessee's Achievement School District and Innovation Zones." *Educational Evaluation and Policy Analysis*, 2017. doi: 10.3102/016237371770.

APPENDIX A

ADMINISTRATIVE DATA

This appendix provides additional details about the administrative data described in Chapter II of the main report.

APS provided all administrative data used in the analyses, which included student assessment, demographic, enrollment, attendance, and suspension data for the 2011–2012 through 2017–2018 school years. The statewide assessment in the 2017–2018 school year was the statewide Georgia Milestones exam, which students in grades 3 through 8 took each spring in ELA, math, science, and social studies. The Georgia Milestones replaced the Criterion-Referenced Competency Tests (CRCT) starting in 2014–2015.[27] The districtwide assessment in 2017–2018 was the STAR exam, which students took multiple times a year, typically in the fall, winter, and spring, in math and ELA or early literacy. The STAR replaced the Computer Adaptive Assessment System (CAAS) exam as the districtwide assessment for elementary grades starting in the 2016–2017 school year. Participation rates in these assessments varied across grades, years, and subjects.[28] Student demographic information included birth month and year, race and ethnicity, gender, English proficiency, disability status, homelessness status, and eligibility for the free/reduced-price lunch program.[29] For each school where a student enrolled, APS also provided the dates of enrollment and attendance and suspension records in every school year.

In addition, APS provided data specifically related to the Turnaround Strategy for the 2017–2018 school year. We obtained roster data for students who worked with CIS site coordinators, High Impact Tutoring (HIT) tutors[30], and math and reading specialists in 2017–2018. Finally, APS provided data on school participation in the various components of the Turnaround Strategy in 2017–2018, as well as schools' designated cluster and Community Eligibility Provision (CEP) status.

We merged the various data sets provided using an anonymized student ID provided by APS and reduced the analytical data set to the student level for the CIS and specialist analyses, and to the student-year level for the targeted and partnership schools' analyses. For students who attended multiple schools within a school year, we counted the student as attending the school where he or she had the greatest number of enrollment days for that school year and summed the total numbers of days suspended, enrolled, and absent across all schools the student attended in the year.

[27] Beginning in the 2016–2017 school year, only students in grades 5 and 8 take the Georgia Milestones science and social studies exams.

[28] In the 2015–2016 school year, the majority of students in kindergarten through grade 11 took the math and reading CAAS exams. In 2016–2017, the majority of students in pre-kindergarten through grade 1 took the STAR early literacy exam, whereas the majority of students in grades 2 through 5 took the STAR ELA exam. Most students in grades 1 through 5 took the STAR math exam.

[29] All students who attended a Community Eligibility Provision (CEP) school qualified for the free/reduced-price lunch program regardless of personal eligibility.

[30] HIT rosters included students from three targeted schools. The math and reading specialist and CIS propensity-score models used indicators of whether students worked with a HIT tutor. See Appendices C and D for more information.

We created several new variables to facilitate the analyses. For example, we transformed student assessment scaled scores into standardized *z*-scores based on districtwide year-, grade-, and subject-specific means and standard deviations. We used suspension records to create indicators of whether students received in- or out-of-school suspensions during the year or after the first quarter of the school year. We also created an indicator of whether students were disciplined for a serious offense during the school year. We determined whether students had been enrolled in the district for only part of each school year and used attendance and enrollment data to calculate students' yearly absence rates, and create an indicator of whether students were chronically absent. We also used students' birth month and year to create an indicator of whether they were behind grade level for their age.

APPENDIX B

TECHNICAL APPENDIX FOR ANALYSIS OF TARGETED AND PARTNERSHIP SCHOOLS

This appendix presents supplemental information for the impact analyses of the Turnaround Strategy's overall targeted supports and school partnerships. We provide additional details about the methodology used and then present results from supplemental analyses.

Supplemental information on the methodology used to evaluate the impacts of overall targeted and partnership supports

We used a quasi-experimental research design known as difference-in-differences to evaluate the impacts of each year of the overall targeted supports and Kindezi and Purpose Built Schools (PBS) school partnerships. To conduct this difference-in-differences analysis, we used student-level data from 2011–2012 to 2017–2018. As discussed in the main text, all 13 targeted schools began receiving support in 2016–2017. The first school partnership—PBS at Thomasville Heights Elementary—also began in 2016–2017. For these schools, we examined impacts after the first and second years of supports. In addition, three other schools began partnerships in 2017–2018: Price Middle and Slater Elementary partnered with PBS, and Gideons Elementary partnered with Kindezi. For these three schools, we were able to examine the impacts of the partnerships after only one year.

Table B.1 lists all of the schools included in the analyses. In selecting comparison schools, we considered only APS (non-charter, traditional) elementary and middle schools that did not participate in the Turnaround Strategy. In addition, we excluded schools in the North Atlanta and Grady clusters from the comparison group because the demographic profile and average academic performance of these schools differed substantially from the schools that participated in the Turnaround Strategy. Based on these selection criteria, we identified 15 comparison schools that operated during the six most recent school years in the analysis. [31]

We measured the impacts of each year of targeted supports or partnerships on six key outcomes: student achievement in ELA, math, science, and social studies, and students' likelihood of being suspended or chronically absent. We measured student achievement by using the state assessment—the CRCT from 2011–2012 to 2013–2014 and the Georgia Milestones from 2014–2015 to 2017–2018. Because scales on state tests changed over time, we converted all scaled scores to normalized *z*-scores within year, grade, and subject. To make the tested grades consistent over time, we examined science and social studies performance among students in grades 5 and 8 only.

[31] Some targeted schools were reconfigured as part of the Turnaround Strategy. In those cases, we reconstructed their baseline period using the schools' prior configuration. In 2016–2017, Hollis Innovation Academy opened and Bethune Elementary School closed. Most of the students in Hollis's first year had gone to Bethune the year before, so we linked those two schools over time. In 2016–2017, Grove Park Intermediate School and Woodson Primary Elementary School merged to form Woodson Park Academy. Most of the students in Woodson Park's first year had gone either to Woodson Primary or Grove Park the year before, so we similarly linked these three schools. Finally, in 2016–2017, Connally Elementary School and Venetian Hills Elementary School merged to form Tuskegee Airmen Global (TAG) Academy. Most of the students in TAG's first year had gone either to Connally or Venetian Hills the year before, so we also linked those three schools.

Table B.1. Schools included in the impact analyses

Targeted schools	Kindezi partnership schools	PBS partnership schools	Comparison schools
Barack and Michelle Obama (BAMO) Elementary	Charles L. Gideons Elementary	Thomasville Heights Elementary	Beecher Elementary
Bazoline E. Usher Elementary		Luther Judson Price Middle	Benteen Elementary
F.L. Stanton Elementary		Thomas Heathe Slater Elementary	Burgess-Peterson Elementary
George A. Towns Elementary			Cleveland Avenue Elementary
Hollis Innovation Academy			Deerwood Academy
L.O. Kimberly Elementary			Emma Hutchinson Elementary
Margaret Fain Elementary			Fred A. Toomer Elementary
T.J. Perkerson Elementary			Harper-Archer Middle
Tuskegee Airmen Global Academy			Heritage Academy Elementary
William Finch Elementary			John Wesley Dobbs Elementary
William J. Scott Elementary			Joseph Humphries Elementary
William M. Boyd Elementary			M. Agnes Jones Elementary
Woodson Park Academy			Martin L. King Jr. Middle
			Parkside Elementary
			West Manor Elementary

PBS = Purpose Built Schools.

To estimate the impacts, we used the following difference-in-differences OLS regression model:

(B1) $y_{igst} = \alpha + \beta T_{igt} + \gamma T_{ig1} + \theta T_{ig2} + \rho post_t + X_{igt}\delta_1 + W_{st}\delta_2 + d_g + d_s + d_t + \varepsilon_{igst}$,

where y_{igst} is the outcome for student i in grade g in school s in year t. T_{igt} indicates whether student i in grade g was enrolled in a targeted (partnership) school in year t and $post_t$ indicates whether year t was in the post-intervention period (which varied across schools in the PBS analysis). T_{ig1} indicates whether student i in grade g was enrolled in a targeted (partnership) school in the first year of supports; similarly, T_{ig2} indicates whether student i in

grade g was enrolled in a targeted (partnership) school in the second year of supports, if relevant. Therefore, γ and θ represent the difference-in-differences estimates, or the impacts of the first and second years of supports, respectively.

The model also accounts for student and school characteristics. X_{igt} is a vector of characteristics for student i in grade g in year t, which includes gender; race/ethnicity; disability status; English language learner status; eligibility for free or reduced-price lunch; and indicators for whether the student was behind grade level for his/her age, homeless, or attended the school for less than the full school year. W_{st} is a vector of characteristics for school s in year t, which includes whether the school was a middle school (rather than an elementary school) and the school's percentage of disabled, free or reduced-price lunch, African American, and English language learner students, as well as whether the school had Community Eligibility Provision (CEP) status. The inclusion of student and school characteristics helps improve the precision of the estimates and account for compositional changes over time. d_g, d_s, and d_t represent grade, school, and year fixed effects, respectively. The school fixed effects account for any remaining differences across schools that are constant over time, whereas the year fixed effects account for aggregate trends. ε_{gsti} is a random error term that reflects the influence of unobserved factors on the outcome. Standard errors are clustered at the school level to account for the lack of independence in student outcomes within schools.

Because PBS schools began receiving supports in different years, we first recentered the annual data around the period when supports began. For Thomasville Heights Elementary, Year 1 of supports corresponded to the 2016–2017 school year and Year 2 corresponded to the 2017–2018 school year. For the other two PBS schools, there was only one year of supports, which corresponded to 2017–2018. Therefore, the Year 2 impact for PBS reflects Thomasville Heights Elementary only. We recentered comparison schools in the same way as Thomasville Heights. The analyses of targeted supports and the Kindezi partnership did not require recentering the data because all participating schools began receiving support in the same year.

Because we evaluated the impacts of targeted supports and PBS and Kindezi partnerships on multiple outcomes and years, the probability that one of those impacts is statistically significant is greater than the probability that a single impact appears statistically significant. We therefore applied the Benjamini-Hochberg multiple comparisons correction to the p-values of each set of impact estimates by year (Benjamini and Hochberg 1995).

To interpret the estimated impacts on test scores, measured in standard deviations, we examined how the academic performance of the average student would change as a result of these impacts. We calculated the average *z*-scores of students in targeted, PBS, and Kindezi schools in the year before supports were implemented and added the Year 1 (and, if applicable, the Year 2) impact estimates for the corresponding subject. We then used the standard normal distribution to determine the corresponding percentile ranks of those *z*-scores.

Results from supplemental analyses

As mentioned in the main text, a key assumption of the difference-in-differences methodology is that all differences between targeted, PBS, or Kindezi schools and comparison

schools were stable over time with the exception of the turnaround supports received. We conducted a series of supplemental analyses to assess potential threats to this assumption.

Compositional changes in students enrolled

Although using different samples of students each year enables us to include more years and students in the analysis, a potential risk is that the composition of students at targeted, PBS, or Kindezi schools could be affected by the turnaround supports either in the year the Turnaround Strategy was announced or after it began. Such compositional changes could result in changes in schoolwide academic performance due to a difference in student body rather than the turnaround supports themselves. For example, this could occur if schools began recruiting higher-performing students than in years past.

To assess this risk, we estimated a version of equation B1 only with students who had outcome data and were enrolled continuously in each type of school (targeted, PBS, Kindezi, or comparison) during the last three school years.[32] We were not able to examine science and social studies performance in this sensitivity analysis because only students in grades 5 and 8 take those tests; therefore, most of the students in this restricted sample did not have any science or social studies scores before supports were implemented.

The results of this analysis appear in Table B.2. Many of the results are similar in sign and magnitude to those of the main analysis. For example, there were consistently positive impacts in math scores, together with increases in reported suspension rates. In addition, the PBS partnership showed changing impacts on chronic absence (a decrease in Year 1 followed by an increase in Year 2), and the Kindezi partnership at Gideons Elementary showed large positive impacts on both ELA and math. It is also worth noting that the main results are not sensitive to the inclusion of student- or school-level covariates (results not shown). Taken together, these sensitivity analyses suggest that compositional changes among students are not driving the results.

[32] We focused on a three-year period because most of the schools being studied were elementary schools; therefore, students did not have test scores for a longer period of time. This fact shortened the pre-intervention period to either one or two years, depending on when supports were implemented. For many schools, the sample comprised just one cohort of students (for example, students in targeted schools who were in grade 3 in 2015–2016 and stayed in a targeted school through 2017–2018).

Table B.2. Impacts based on a stable set of students

	ELA	Math	Likelihood of being suspended	Likelihood of being chronically absent
Impact of targeted supports in Year 1	-0.01	0.06	0.01	0.01
	(0.05)	(0.05)	(0.01)	(0.01)
Impact of targeted supports in Year 2	-0.05	0.10	0.06**	0.02
	(0.06)	(0.07)	(0.02)	(0.02)
Impact of PBS in Year 1	0.05	0.24*	0.10**	-0.04**
	(0.07)	(0.10)	(0.02)	(0.01)
Impact of PBS in Year 2	-0.13*	0.18**	0.08**	0.11**
	(0.05)	(0.06)	(0.01)	(0.01)
Impact of Kindezi in Year 1	0.33**	0.40**	0.17**	-0.00
	(0.04)	(0.05)	(0.01)	(0.01)
Number of students in analysis	8,339	8,256	31,275	31,275
Number of schools in analysis	32	32	32	32

Source: APS administrative data.

Note: This table displays impact estimates for the 2018 Georgia Milestones end-of-grade exams in *z*-scores (standard deviations) and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the likelihood that a student was ever suspended during the school year. "Chronically absent" refers to the likelihood that a student was missing 10 percent or more of days enrolled. Standard errors are displayed in parentheses below each impact estimate.

**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; ELA = English language arts.

Anticipatory effects

With any school turnaround effort, it is possible that the announcement of changes to come and the uncertainty and disruption that can follow could lead to anticipatory effects in the school year before supports were implemented. For example, anticipation of the partnerships with PBS or Kindezi could have led to staff turnover before the partnerships began, causing academic performance in the school to drop. In this scenario, improvements in the first year of the partnerships could be the result of a natural rebound rather than the partnerships. Conversely, if anticipation of the partnerships led to positive changes in the school before supports began, the difference-in-differences analysis could understate the improvements in the first year.

Although schools' academic performance trends did not suggest that anticipatory effects occurred (Figures III.2, VI.1, and VI.3), we conducted a sensitivity analysis that excluded data from the year before supports were implemented and compared these results with the main findings based on all years. If anticipation of the partnership affected student outcomes in the year before supports were implemented, the results could change. However, the results of this sensitivity analysis are consistent with the main findings in both sign and magnitude, suggesting that there were limited anticipatory responses in these schools (Table B.3).

Table B.3. Impacts of targeted supports and partnerships when the year before supports began is excluded from the analysis

	ELA	Math	Science	Social studies	Likelihood of being suspended	Likelihood of being chronically absent
Impact of targeted supports in Year 1	-0.03	0.01	-0.15	-0.15	-0.00	0.03
	(0.04)	(0.05)	(0.08)	(0.08)	(0.02)	(0.01)
Impact of targeted supports in Year 2	-0.02	0.11**	-0.08	-0.05	0.03	0.05**
	(0.03)	(0.04)	(0.09)	(0.08)	(0.02)	(0.02)
Impact of PBS in Year 1	-0.04	0.10	-0.13	-0.20	0.09**	-0.02**
	(0.04)	(0.05)	(0.17)	(0.19)	(0.01)	(0.01)
Impact of PBS in Year 2	-0.11**	0.23**	-0.19*	-0.06	0.05**	0.09**
	(0.03)	(0.04)	(0.08)	(0.07)	(0.01)	(0.01)
Impact of Kindezi in Year 1	0.11**	0.41**	-0.39**	-0.49**	0.14**	-0.03**
	(0.03)	(0.03)	(0.07)	(0.06)	(0.01)	(0.01)
Number of students in analysis	42,757	42,704	13,378	13,424	91,639	91,639
Number of schools in analysis	32	32	32	32	32	32

Source: APS administrative data.

Note: This table displays impact estimates for the 2018 Georgia Milestones end-of-grade exams in *z*-scores (standard deviations) and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the likelihood that a student was ever suspended during the school year. "Chronically absent" refers to the likelihood that a student was missing 10 percent or more of days enrolled. Standard errors are displayed in parentheses below each impact estimate.

**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; ELA = English language arts; PBS = Purpose Built Schools.

APPENDIX C

TECHNICAL APPENDIX FOR ANALYSIS OF MATH AND READING SPECIALISTS

This appendix presents supplemental information for the analysis of math and reading specialists. We first present supplemental information on the specialist roster data by school. We then present additional information on the methodology used to evaluate the intervention. Finally, we present supplemental results for the impact analyses presented in the main text.

Supplemental information on math and reading specialist data

We received a roster of students who received specialist services from APS. The roster included the subject of those services (math or reading) and the start dates of services for most students. We imputed missing start dates with the average nonmissing start date for that subject within the student's school.[33] Only one school, Barack and Michelle Obama (BAMO) Elementary, recorded end dates, and only for a subset of students. However, because students were supposed to receive support through the end of the school year, we imputed missing end dates per APS's guidance, using the last day each student was enrolled in the school where they received support. Across all targeted schools, 464 students were included in the roster as working with a math specialist and 460 with a reading specialist. Among these students, 101 worked with both a math and reading specialist in the 2017–2018 school year. Table C.1 includes the number of math and reading specialist students for each school.

Table C.1. Math and reading specialist enrollment, by school

School	Total school enrollment	Math specialist only	Reading specialist only	Both math and reading specialists
BAMO Elementary School	304	41	–	13
Boyd Elementary School	544	34	50	13
F.L. Stanton Elementary School	348	51	59	–
Fain Elementary School	529	20	44	–
Finch Elementary School	606	32	20	–
Hollis Innovation Academy	652	29	28	–
Kimberly Elementary School	516	–	–	64
Perkerson Elementary School	505	28	12	–
Scott Elementary School	481	–	20	–
Towns Elementary School	427	28	34	5
Tuskegee Airmen Global Academy	784	37	38	–
Usher/Collier Elementary School	505	33	26	–
Woodson Park Academy	712	30	24	–
All targeted schools	**6,913**	**363**	**359**	**101**

Source: APS administrative data.

Notes: To protect students' identities, we replaced all cells where less than five students were enrolled with a "–".

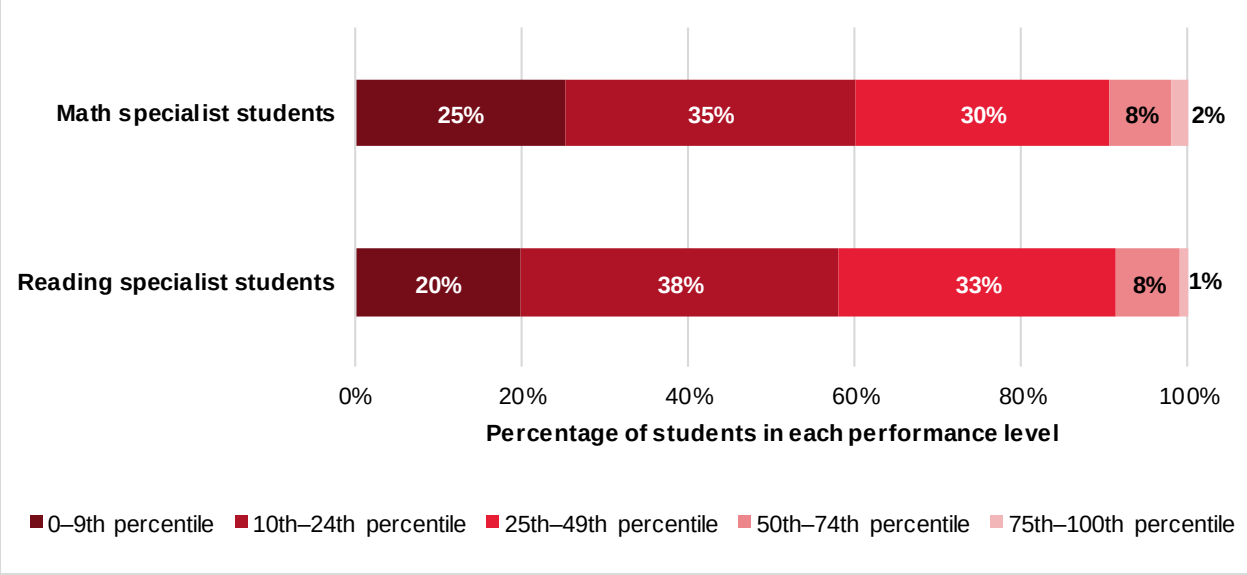
APS = Atlanta Public Schools.

As mentioned in the main text, math and reading specialists were expected to work with students in the bottom 5 to 10 percent of academic performance at the school. However, Figure IV.5 showed that specialists worked with students across a variety of achievement levels (based

[33] In one school, Woodson Park Academy, all math specialist start dates were missing. For students who also worked with a reading specialist, we imputed the missing math start date with their reading start date. Otherwise, we imputed the missing math start date with the school's average reading start date. In all cases where we imputed a missing start date, we verified that the date did not fall before the student's enrollment date in the school from which the student received specialist services.

on fall 2017 STAR scores) relative to their peers in the same school. In Figure C.1, we compare the baseline performance of specialist students relative to other students in the district, rather than their same school, since many students in targeted schools had low academic performance. When compared to the rest of the district, a higher percentage of specialist students scored in the bottom 10th percentile on the fall STAR exams than they did in their respective schools. In other words, more of the students working with specialists in targeted schools were in the bottom decile of the district-wide distribution than were in the bottom decile of their school's distribution. However, even compared to the rest of the district, specialist students' achievement levels varied at baseline, with about 40 percent in each subject scoring above the 50th percentile districtwide.

Figure C.1. Performance of math and reading specialist students on the fall 2017 STAR assessment relative to other students in APS district



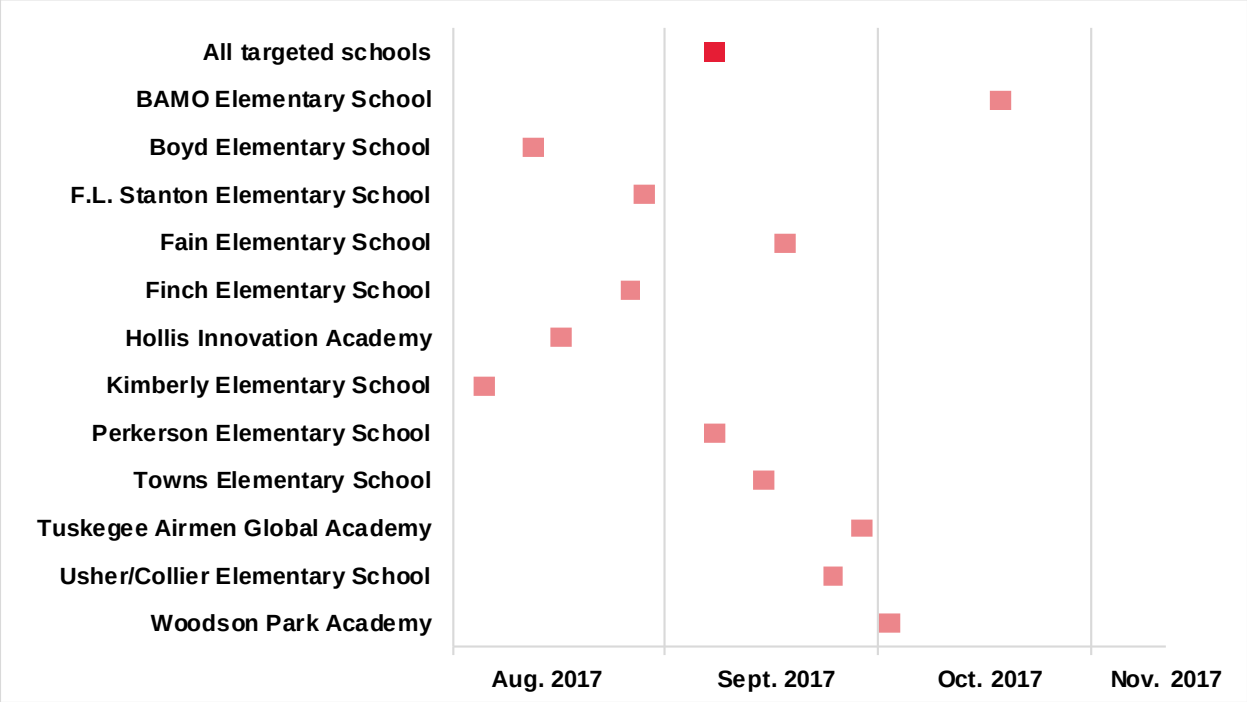
Source: APS administrative data.

Notes: Figure reads: Thirty percent of students who worked with math specialists scored in the 25th to 49th percentile of students across the district on the fall 2017 STAR assessment.

APS = Atlanta Public Schools.

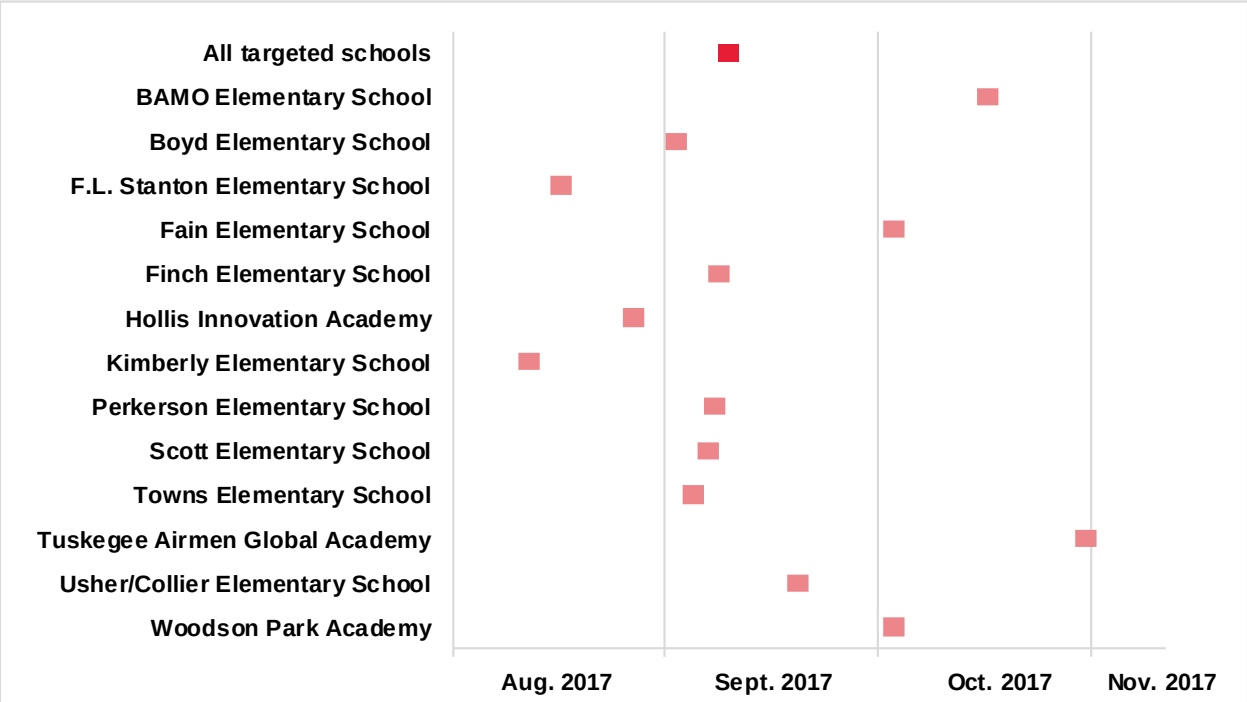
Figures C.2 and C.3 show the date that the average student was listed as entering the math or reading specialist roster, respectively, for each school. The typical student who worked with a math or reading specialist began receiving services in early September, although there was some variation across schools. For example, all math and reading specialist students at Kimberly Elementary School entered the roster in early August; whereas the average math and reading specialist students at BAMO Elementary School began receiving services in mid-October. No other information on the intensity or frequency of specialist services was available.

Figure C.2. Average start date of math specialist support, by school



Source: APS administrative data.
APS = Atlanta Public Schools.

Figure C.3. Average start date of reading specialist support, by school



Source: APS administrative data.
APS = Atlanta Public School.

Supplemental information on the methodology used to evaluate math and reading specialists

Sample selection

The math and reading specialist impact analyses included students who met the following criteria: (1) were in grades 1 through 5;[34] (2) received specialist services for 10 weeks or more; (3) had baseline performance data (fall 2017 STAR scores) and demographic data; (4) were successfully matched to similar comparison students; and (5) had outcome scores in spring 2018 for the respective subject (either on the Georgia Milestones or STAR assessments). The eligible comparison groups were composed of 1st through 5th grade students who were also enrolled in a targeted school and had baseline data available but did not receive specialist services for the respective subject. Table C.2 summarizes how each of the above restrictions affected the sample sizes for the analyses.

Table C.2. Summary of math and reading specialist sample sizes

	Reading	Math
Total number of specialist students	460	464
Specialist students in 1st–5th grade	404	419
Specialist students with at least 10 weeks of services	403	414
Specialist students with baseline data	302	335
Specialist students matched	300	318
Comparison students matched	1,610	1,737
Matched specialist students with 2018 Milestones scores	211	240
Matched specialist students with 2018 STAR scores	291	316

Source: APS administrative data.

Notes: "Specialist students matched" refers to specialist students who had baseline data and matched with at least one comparison student.

APS = Atlanta Public Schools.

To be included in the specialist analyses, students had to have STAR math and STAR ELA or STAR early literacy (SEL) scores from fall 2017 as well as other baseline data (see Table C.3 for a complete list of baseline variables). We also matched students on STAR math and STAR ELA or SEL scores from the 2016–2017 school year. However, some students were missing these scores so this information was not required for a student to be included in the analysis.

[34] In the 2017–2018 school year, 42 kindergarten students worked with a math specialist and 54 worked with a reading specialist. Kindergarteners were not included in the analyses because baseline data are not available. Sixth grade students were also omitted from the analyses due to an insufficient sample size.

Instead, we imputed missing values using a dummy imputation method in which, for each exam, we created an indicator of whether the student was missing that exam and then set the score to a constant value of zero.[35] We used the STAR exam rather than Georgia Milestones to account for prior academic performance because students in all elementary grades took these exams, enabling us to include grades 1 through 5 in the analysis.

Students who worked with a specialist and met the sample selection criteria above were eligible to be included in the analysis, even if they attended another school for a greater number of days in the 2017–2018 school year. About 2 percent of math and less than 1 percent (0.25 percent) of ELA specialist students were enrolled at another school for more days in the 2017–2018 school year.

Propensity-score matching methodology

For each subject of specialist support, we estimated a propensity score for each eligible specialist and comparison student by grade, using a logistic regression model. This propensity score indicates the likelihood of receiving support from a specialist in that subject, given students' prior academic performance and other characteristics. Table C.3 lists the variables used to estimate the propensity scores. As mentioned previously, we accounted for baseline academic performance by using STAR scores from fall 2017 and the previous school year. For exams taken in the 2016–2017 school year, we used the student's most recent available score from the spring, winter, or fall testing windows.[36] For STAR, students could take the ELA or SEL tests. If available, we used students' ELA scores. Otherwise, we used their SEL scores, which were more common among students in early grades.[37] We first standardized all test scores by school year, grade, and subject.

After generating the propensity scores, we matched each eligible specialist student with up to 20 comparison students who had the most similar propensity scores within a given threshold or radius of the specialist student's propensity score.[38] If there were no eligible comparison students within the matching radius for a given specialist student, that student was excluded from the matched comparison impact analyses. As summarized in Table C.2, we were able to match 318 out of 335 math specialist students with baseline data and 300 out of 302 reading specialist

[35] Before matching, 40 percent of specialist and potential comparison students were missing a STAR math score from the previous year and 21 percent were missing an ELA or SEL STAR score. Among students matched, about 14 percent were missing fall 2017 STAR scores; less than 1 percent were missing fall ELA or SEL STAR scores. We included the imputed test scores and missing value indicators in the propensity score estimation and impact analyses.

[36] For the 2016–2017 STAR ELA exam, 95 percent of scores used were from the spring, 3 percent from the winter, and 2 percent from the fall. For the STAR math exam, 89 percent of scores used were from the spring, 6 percent from the winter, and 5 percent from the fall.

[37] About 5 percent of reading specialist and matched students had an SEL rather than ELA score as their fall 2017 STAR baseline score; 29 percent had an SEL rather than ELA score as a 2016–2017 STAR baseline score.

[38] The same matching radii were used for the math and reading specialist analyses. The radii ranged from 0.04 to 0.325 for each grade. The radii for each grade were selected to improve the quality of matches obtained.

students. On average, each matched math and reading specialist student was paired with 19 comparison students.

Table C.3. Baseline variables used in the math and reading specialist propensity-score models

Math and ELA or SEL STAR scores from fall 2017
Math and ELA or SEL STAR scores from the 2016–2017 school year
Second and third order math and ELA or SEL STAR scores from fall 2016
Second and third order math and ELA or SEL STAR scores from the 2016–2017 school year
Indicator of whether the student took the ELA or SEL test in fall 2017
Indicator of whether the student took the ELA or SEL test in the 2016–2017 school year
Indicators of whether the 2016–2017 STAR baselines scores were imputed using dummy imputation
Student demographics from the 2016–2017 school year (gender, race/ethnicity, English language learner status, disability status, homelessness status)
Indicator of whether the student was enrolled in a school for only part of the 2016–2017 school year
Indicator of whether the student was suspended at any point in the 2016–2017 school year
Attendance rate for the 2016–2017 school year
Indicator of whether the student was chronically absent in the 2016–2017 school year
Indicator of whether the student was behind grade level for the student's age in the 2016–2017 school year
Interactions of baseline STAR scores and demographic variables (gender, disability status)
Indicator of whether the student worked with a HIT tutor for the respective subject in which specialist services were received
Indicator of whether the student received case management services from CIS during the 2017–2018 school year

Note: A logistic regression was run for each grade and subject. Some variables or interactions listed were omitted from grade-specific models if there was no variation in specialist students for that model. For example, no math specialist students in grade 1 had a disability, so that variable (and corresponding interactions) was omitted from the grade 1 model.

CIS = Community in Schools; ELA = English language arts; HIT = High Impact Tutoring; SEL = STAR early literacy.

Table C.4 presents summary statistics showing how well specialist students were matched to comparison students on baseline characteristics, by subject. On average, comparison students from each matched group were not significantly different from the specialist students on any baseline characteristics used in the analyses. Similarly, there were no statistically significant baseline differences between specialist and matched comparison students in the analytical samples used to estimate impacts for any outcome.

Table C.4. Baseline characteristics of matched specialist and comparison students

Baseline characteristics	Math specialist students	Matched comparison students	Reading specialist students	Matched comparison students
2016–2017 STAR math score (mean *z*-score)	-0.662 (0.692)	-0.644 (0.702)	-0.552 (0.702)	-0.532 (0.678)
2016–2017 STAR ELA or SEL score (mean *z*-score)	-0.647 (0.567)	-0.659 (0.570)	-0.745 (0.576)	-0.740 (0.572)
Fall 2017 STAR fall math score (mean *z*-score)	-0.848 (0.701)	-0.832 (0.689)	-0.710 (0.721)	-0.695 (0.709)
Fall 2017 STAR ELA or SEL reading score (mean *z*-score)	-0.712 (0.481)	-0.727 (0.477)	-0.823 (0.406)	-0.821 (0.410)
Disability	0.025	0.029	0.027	0.026
English language learner	0.019	0.019	0.023	0.023
Homeless	0.041	0.031	0.017	0.014
Partial enrollment	0.248	0.241	0.207	0.207
Ever suspended	0.025	0.026	0.050	0.047
Attendance rate	0.941 (0.049)	0.940 (0.058)	0.944 (0.050)	0.946 (0.049)
Chronically absent	0.167	0.168	0.167	0.158
Behind grade level	0.072	0.068	0.077	0.084
Hispanic	0.022	0.021	0.030	0.029
American Indian	0.016	0.013	0.023	0.021
Asian	0.003	0.005	0.000	0.003
African American	0.978	0.977	0.973	0.976
Pacific Islander	0.000	0.002	0.000	0.001
Female	0.582	0.578	0.537	0.539
HIT participant (in respective subject)	0.022	0.026	0.030	0.030
CIS participant	0.104	0.102	0.083	0.080
Number of students	**318**	**1,737**	**300**	**1,610**

Source: APS administrative data.

Notes: The math and reading specialist and their respective matched comparison groups did not significantly differ on any variables. Standard deviations are displayed in parenthesis below the averages of continuous variables.

APS = Atlanta Public Schools; CIS = Community in Schools; ELA = English language arts; HIT = High Impact; Tutoring; SEL = STAR early literacy.

Impact model

To measure impacts using the matched samples for each subject, we estimated an ordinary least squares (OLS) regression model that accounted for any small remaining differences between specialist and comparison students in their prior academic performance and characteristics:

(C1) $$y_i = \alpha + X_i\beta + \delta T_i + grade\ dummies + school\ dummies + \varepsilon_i$$

where y_i is the outcome of interest for student i; X_i is a vector of demographic controls and baseline test scores for student i; T_i is a binary variable for treatment status, indicating whether student i received support from a specialist in a given subject; ε_i is a random error term that reflects the influence of unobserved factors on the outcome; and δ and β are parameters or vectors of parameters to be estimated, with δ representing the impact of intervention of interest. Because each comparison student could be matched to multiple specialist students, we used a weighting scheme in which each student served by a specialist had a weight of one, and each comparison student had a weight representing the fraction of the number of matching specialist students.

We assessed the impact of math and reading specialists on the Georgia Milestones exam taken by students in grades 3 through 5 and on the spring STAR exams for students in grades 1 through 5.[39, 40] Compared to the analyses of Georgia Milestones scores, those of student performance on STAR included more students (see Table C.2) because the assessments are used in more grades. Because we evaluate the effect of specialists on two exams in the same subject, the probability that one of those two impacts is statistically significant is greater than the probability that a single impact appears statistically significant. To compensate for the number of inferences being made across exams within a given subject, we applied the Benjamini-Hochberg multiple comparisons correction to the p-values of each pair of impact estimates by subject (Benjamini and Hochberg 1995).

To interpret the estimated impacts on test scores, measured in standard deviations, we examined how the academic performance of the average student would change as a result of these impacts. For each subject, we calculated the average z-scores of specialist students on the fall 2017 STAR assessment and added the impact estimates for the corresponding subject. We then used the standard normal distribution to determine the corresponding percentile ranks of those z-scores.

[39] We also assessed the impact of math and reading specialists on the spring STAR exams in grades 3 through 5 only. No meaningful differences were found between these results and the results from the Milestones analysis in either specialist group. (Results not shown.)

[40] Similar to the baseline STAR exams, we also used spring SEL scores for the ELA outcome when the ELA outcome was not available. About 2 percent of reading specialist and matched students had an SEL rather than ELA score as their spring 2018 STAR outcome score.

Supplemental results on analyses presented in the main text

In addition to the primary impact analyses summarized in the main text, we conducted exploratory analyses to assess whether the impacts differed for specific groups of students. Specifically, we tested whether the impacts of specialists differed (1) between students in schools where their specialist provided support for a longer period of time (about 7.5 months or more, compared to less time),[41] and (2) between students who were among the lowest performing in their school on the fall STAR exams and those who had relatively higher performance.[42]

For each subject, we estimated the following regression model, which adds interaction terms to the benchmark model in equation C1.

(C2) $$y_i = \alpha + X_i\beta + \delta T_i + \gamma T_i E_i + grade\ dummies + school\ dummies + \varepsilon_i$$

The coefficient γ represents how the impact differs for the exploratory variable of interest (E_i) (for example, whether students were among the lowest performing in their school). Because students were not randomly assigned to a duration of support, nor is prior academic performance random, these analyses are exploratory and might reflect the influence of other related but unobserved factors.

The impact of specialists on assessment scores by duration of support

Table C.5 presents the impacts of math and reading specialists based on the average number of months specialists provided support in their schools. Math and reading specialists in three schools (Woodson Park Academy, TAG Academy, and BAMO Elementary School) were identified as having provided support for a shorter period of time in their respective subject compared to other targeted schools. We found no statistical evidence to support the notion that more time spent providing support made a difference in student achievement in either subject or assessment.

[41] In a quarter of schools, students received math support for less than 7.7 month and reading support for less than 7.4 months on average, so we used these values to classify schools as longer or shorter duration. Average duration varied across schools primarily because of different start dates and differences in student mobility (the majority of end dates were imputed using the date of each student's enrollment in the school). Specialists were generally expected to work with students for all or most of the school year.

[42] Low-performing students were defined at those who scored in the bottom 25th percentile of the school in which they received specialist services on the fall 2017 STAR exam in the respective subject.

Table C.5. Impacts of math and reading specialists on assessment scores, by schoolwide average specialist support length

	Math		Reading	
	Milestones Math	STAR Math	Milestones ELA	STAR ELA
Impact of working with a shorter-serving specialist	0.06	0.08	0.06	0.01
	(0.08)	(0.09)	(0.08)	(0.08)
Impact of working with a longer-serving specialist	0.06	0.07	0.07	0.05
	(0.04)	(0.05)	(0.05)	(0.03)
Difference in impacts between longer- and shorter-serving specialists	0.00	-0.01	0.01	0.05
	(0.09)	(0.10)	(0.09)	(0.09)
Number of students	**1,317**	**1,579**	**1,347**	**1,636**

Source: APS administrative data.

Notes: This table displays impact estimates in *z*-scores (standard deviations) on the 2018 Georgia Milestones exams and spring 2018 STAR exams for the same subject as the specialist services. "Shorter-serving specialists" are those in schools where average duration of support was in the bottom quartile compared to other targeted schools. "Longer-serving specialists" are those in schools with average duration above the bottom quartile. The Georgia Milestones estimates are based only on students in grades 3 through 5. Standard errors are displayed in parentheses below each impact estimate. The sample size reflects the total number of specialist students and matched comparison students in each analysis.

**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools; ELA = English language arts.

The impact of specialists on assessment scores by baseline academic performance

We tested the hypothesis that students who were lower performing at baseline could have benefited more from working with a math or reading specialist. Thirty-four percent of students in grades 1 through 5 who worked with a reading specialist were in the bottom quartile in their school based on the fall 2017 STAR ELA exam. In math, 43 percent of students who worked with a specialist were in the bottom quartile in their school. Reading specialists had a statistically significant impact of 0.12 standard deviations on Milestones ELA scores among students who scored in the bottom 25th percentile at baseline (Table C.6). Reading specialists' impacts on STAR scores, and math specialists' impacts on both Georgia Milestones and STAR scores, were also greater among these lower performing students, although the differences were not statistically significant.

Table C.6. Impacts of math and reading specialists on assessment scores, by baseline academic performance

	Math		Reading	
	Milestones Math	STAR Math	Milestones ELA	STAR ELA
Impact of specialist support on higher-performing students	0.05 (0.04)	0.05 (0.05)	0.04 (0.05)	0.04 (0.03)
Impact of specialist support on lower-performing students	0.08 (0.05)	0.12 (0.06)	0.12* (0.05)	0.05 (0.04)
Difference in impacts between lower- and higher-performing students	0.03 (0.06)	0.07 (0.08)	0.08 (0.06)	0.01 (0.05)
Number of students	**1,418**	**1,702**	**1,347**	**1,636**

Source: APS administrative data.

Notes: This table displays impact estimates in *z*-scores (standard deviations) on the 2018 Georgia Milestones exams and spring 2018 STAR exams for the same subject as the specialist services. "Lower-performing students" are those who scored in the bottom quartile on the fall 2017 STAR exam of the respective subject compared to other students in their school. "Higher-performing students" are those who scored above the bottom quartile. The Georgia Milestones estimates are based only on students in grades 3 through 5. Standard errors are displayed in parentheses below each impact estimate. The sample size reflects the total number of specialist students and matched comparison students in each analysis.

**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

APS = Atlanta Public Schools.

APPENDIX D

TECHNICAL APPENDIX FOR ANALYSIS OF COMMUNITIES IN SCHOOLS (CIS) CASE MANAGEMENT

This appendix presents supplemental information for the Communities in Schools (CIS) analyses. We first present supplemental information on CIS data by school. We then present additional information on the methodology used to evaluate the intervention. Finally, we present supplemental results for the impact analyses presented in the main text.

Supplemental information on CIS data

APS provided CIS case management rosters for each month of the 2017–2018 school year and an activity log containing all of the activities documented by the CIS site coordinators during the year for each student with whom they worked. The rosters included the student's grade and school in which they received CIS services. The activity log included the activity category; whether the activity was at the individual, small group, or school level;[43] the activity duration (in minutes); and the activity date. We used the activity log data to create several new student-level variables for the impact analyses. We used the first and last activity dates to calculate each student's duration of CIS support. We also calculated the total time spent on activities (in hours) by summing the activity durations for each student and the total number of activities (overall and by tier) by summing the number of activities logged.

We combined the monthly roster files into one file for the school year. The combined roster contained 357 students, 20 of whom did not have activity data. Conversely, the activity data contained 395 students, 58 of whom were not included in the roster. Per guidance from APS, we based our analysis only on students who had activity data recorded (even if they did not appear in the roster) because the activity log included those students who actually received CIS services.[44] Table D.1 lists the number of CIS students in each targeted school in addition to a monthly summary of the service data logged per CIS student. Differences across schools could reflect actual differences in the services provided as well as differences in data entry practices among site coordinators.

[43] Activity tiers included schoolwide services (Tier I), targeted small group support (Tier II), and individual case management (Tier III). Seventy-four percent of students logged a Tier I activity, 95 percent logged a Tier II activity, and 91 percent logged a Tier III activity. Although 91 percent of students received at least some Tier III support, they experienced more Tier II activities on average.

[44] CIS staff noted that some CIS students received very limited support but could have been included in the service data without appearing in the roster. Thus, we excluded students with fewer than five individual or small group activities logged from the impact analysis.

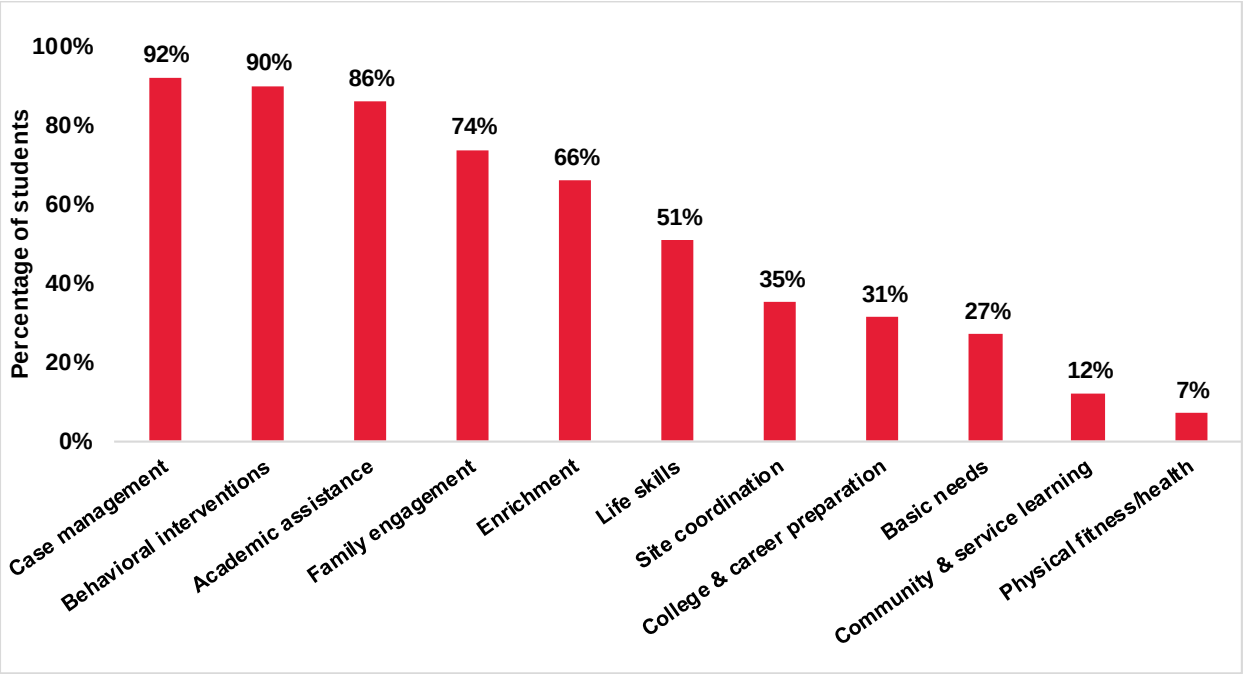
Table D.1. CIS enrollment and services, by school

School	CIS students	Average number of logged activities per student	Average number of logged hours per student	Average monthly number of logged activities per student
BAMO Elementary School	26	37.9	43.3	5.3
Boyd Elementary School	24	16.3	15.2	3.0
F.L. Stanton Elementary School	34	12.6	7.1	2.8
Fain Elementary School	30	56.1	36.2	7.2
Finch Elementary School	22	13.0	5.5	5.4
Hollis Innovation Academy	84	13.9	6.3	4.8
Kimberly Elementary School	21	22.0	15.6	3.2
Perkerson Elementary School	25	44.2	34.3	9.9
Scott Elementary School	22	20.2	14.2	2.9
Towns Elementary School	27	60.5	40.5	7.6
Tuskegee Airmen Global Academy	29	48.4	51.4	8.3
Usher/Collier Elementary School	26	20.0	18.6	2.7
Woodson Park Academy	25	24.6	28.8	3.2
All targeted schools	**395**	**28.2**	**22.1**	**5.1**

Figure D.1 shows the percentage of students who had each type of activity logged. Almost all students received case management, behavioral interventions, and academic assistance. Other types of activities that CIS logged were less common. For example, only 27 percent of students received basic needs support, which includes assistance with food, school supplies, and emergency funds for family needs such as utility bills.[45] The majority (about 60 percent) of students logged activities for at least six months, and half logged them for at least seven months. However, as shown in Figure D.2, the duration of services provided to students varied across schools. The average duration between the first and last activity logged ranged from three months in Hollis Academy to 7.7 months in Towns Elementary School.

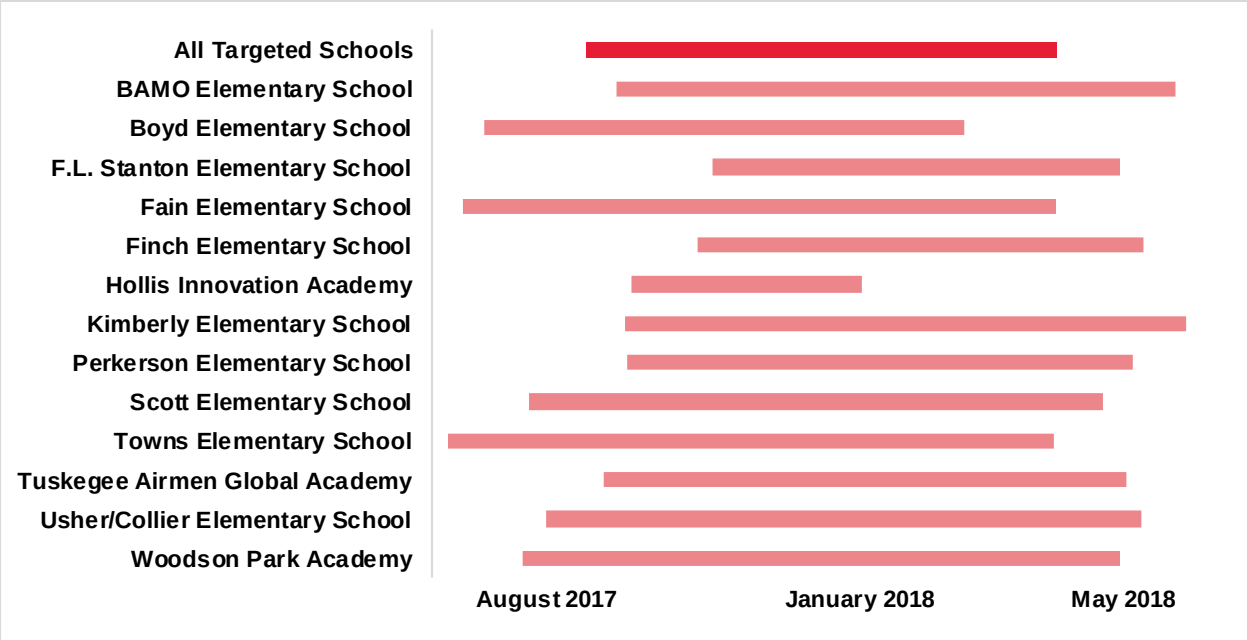
[45] It is important to note some caveats about the activity data. First, some activities, such as case management and behavioral interventions, did not involve actual time spent with the student, but rather checking on attendance rosters and site coordination work. Second, the data are subject to the accuracy and consistency of site coordinators' data entry practices. For example, different site coordinators could have classified similar activities differently.

Figure D.1. Percentage of CIS students experiencing each activity type



Source: APS administrative data.
CIS = Communities in Schools.

Figure D.2. Average CIS duration by targeted school



Source: APS administrative data.
CIS = Communities in Schools.

Supplemental information on the methodology used to evaluate CIS

Sample selection

The CIS impact analyses included students who met the following criteria: (1) were in grades 3 through 5;[46] (2) logged five or more individual or small group activities during the school year; (3) had the required baseline data (described in the next paragraph); (4) were successfully matched to similar comparison students; (5) had math or ELA outcome scores in spring 2018 (either on the Georgia Milestones or STAR assessments), or 2017–2018 attendance and suspension outcome data. The eligible comparison groups were composed of 3rd through 5th grade students who were also enrolled in a targeted school and had baseline data available but did not participate in CIS. Table D.2 summarizes how each of the above restrictions affected the sample sizes for the analyses.

Table D.2. Summary of CIS sample size

	Number of students
Total number of CIS students	395
CIS students in 3rd–5th grade	377
Specialist students with at least five individual or small group activities logged	348
CIS students with baseline data	310
CIS students matched	292
Comparison students matched	1,225
Matched CIS students with 2018 Milestones math scores	288
Matched CIS students with 2018 Milestones ELA scores	288
Matched CIS students with 2018 STAR math scores	247
Matched CIS students with 2018 STAR ELA or SEL scores	250
Matched CIS students with 2018 attendance and suspension data	292

Source: APS administrative data.

Notes: Number of CIS students matched refers to CIS students who had baseline data and matched with at least one comparison student.

CIS = Community in Schools; ELA = English language arts; SEL = STAR early literacy.

To be included in the CIS analysis, students had to have one STAR math score and one STAR ELA or SEL score from the 2016–2017 school year[47] as well as non-academic data from the 2015–2016 and 2016–2017 school years (see Table D.3 for a complete list of baseline variables). We used two years of baseline data in this analysis because students could have been selected to work with CIS at various times between the 2015–2016 and 2017–2018 school years,

[46] In the 2017–2018 school year, there were eight CIS students outside of grades 3 through 5. Because we conducted propensity-score matching at the grade level and a very small number of students in grades pre-K, 1, 2, and 6 participated, they were excluded from the analysis.

[47] Alternatively, we could have required that students have STAR scores from fall 2017, as in the analysis of math and reading specialists. However, unlike specialist students, more CIS students had STAR scores from the 2016–2017 school year than from fall 2017.

and 2015–2016 academic and behavior data proved predictive of CIS participation in the 2017–2018 school year. Although we required test scores only from the 2016–2017 school year, we also matched students on STAR math and STAR ELA or SEL scores from fall 2017 and CAAS scores from the 2015–2016 school year. However, some CIS and comparison students were missing these scores, so this information was not required for a student to be included in the analysis. To keep them in the analysis, we imputed missing values using a dummy imputation method, in which for each exam we created an indicator of whether the student was missing that exam and then set the score to a constant value of 0.[48]

CIS students who met the sample selection criteria above were eligible to be included in the analysis even if they attended another school for a greater number of days in the 2017–2018 school year. About 1.4 percent of CIS students were enrolled at another school for more days in the 2017–2018 school year.

Propensity-score matching methodology

We estimated a propensity score for each eligible CIS and comparison student, by grade, using a logistic regression model. This propensity score indicates the likelihood of participating in CIS case management, given students' prior academic performance and characteristics. Table D.3 lists the variables used to estimate the propensity scores. As mentioned previously, we used baseline data spanning the 2015–2016 school year through fall 2017, just before students began to receive case management in the most recent school year. We accounted for baseline academic performance by using STAR scores from fall 2017 and the previous school year. For exams taken in the 2016–2017 school year, we used the student's most recent available score from the spring, winter, or fall testing windows.[49] For the 2015–2016 CAAS exam, we used the most recent available score from the spring or fall windows.[50] All test scores were first standardized by school year, grade, and subject.

[48] Before matching, 26 percent of CIS and potential comparison students were missing a fall 2017 STAR math score, 24 percent were missing a fall STAR ELA score, and 27 percent were missing CAAS math and reading scores. Among students matched, we imputed about 10 percent of fall 2017 STAR scores, 7 percent of fall STAR ELA scores, and 1 percent of CAAS math and reading scores. The imputed test scores and missing value indicators were included in the propensity-score estimation and impact analyses.

[49] For the 2016–2017 STAR ELA exam, 95 percent of scores used were from the spring, 3 percent from the winter, and 2 percent from the fall. For the STAR math exam, 89 percent of scores used were from the spring, 6 percent from the winter, and 5 percent from the fall.

[50] For both the 2015–2016 CAAS math and reading exams, 96 percent of scores used were from the spring and 4 percent from the fall.

Table D.3. Baseline variables used in the CIS propensity-score models

Math and ELA or SEL STAR scores from fall 2017
Math and ELA or SEL STAR scores from the 2016–2017 school year
Second and third order math and ELA or SEL STAR scores from fall 2016
Second and third order math and ELA or SEL STAR scores from the 2016–2017 school year
Indicator of whether the student took the ELA or SEL test in fall 2017
Indicator of whether the student took the ELA or SEL test in the 2016–2017 school year
Math and reading CAAS scores from the 2015–2016 school year
Indicators of whether CAAS or fall 2017 STAR baselines scores were imputed using dummy imputation
Student demographics from the 2015–2016 and 2016–2017 school years (gender, race/ethnicity, English language learner status, disability status, homeless status)
Indicators of whether the student was enrolled in a school for only part of the 2015–2016 and 2016–2017 school years
Indicators of whether the student was suspended at any point in the 2015–2016 and 2016–2017 school years
Attendance rates for the 2015–2016 and 2016–2017 school years
Indicators of whether the student was chronically absent in the 2015–2016 and 2016–2017 school years
Indicators of whether the student was behind grade level for his/her age in the 2015–2016 and 2016–2017 school years
Indicators of whether the student was suspended for two or more days in the 2015–2016 and 2016–2017 school years
Indicators of whether the student was disciplined at any point in the 2015–2016 and 2016–2017 school years
Indicators of whether the student was disciplined for a serious offense at any point in the 2015–2016 and 2016–2017 school years
Indicator of whether the student was suspended in fall 2017
Indicator of whether the student was disciplined in fall 2017
Indicator of whether the student received academic support in the 2016–2017 school year (HIT or specialist support in either math or reading)
Interactions of baseline STAR scores and 2016–2017 demographic variables (gender, disability status, ever suspended, ever disciplined, ever disciplined for a serious offense, disciplined in fall 2017, chronic absence, academic support)
Interactions of gender and 2016–2017 demographic variables (disciplined in fall 2017, suspended in fall 2017, chronic absence)
Interactions of academic support and 2016–2017 demographic variables (disciplined in fall 2017, suspended in fall 2017, chronic absence, behind grade level)

Note: We ran a logistic regression for each grade and subject. We omitted some variables or interactions listed from grade-specific models if there was no variation in CIS students for that model.

CAAS = Computer Adaptive Assessment System; ELA = English language arts; HIT = High Impact Tutoring; SEL = STAR early literacy.

After generating the propensity scores, we matched each eligible CIS student with up to 20 comparison students who had the most similar propensity scores within a given threshold or radius of the CIS student's propensity score.[51] If there were no eligible comparison students within the matching radius for a given CIS student, we excluded that student from the matched comparison impact analyses. As summarized in Table D.2, we were able to match 292 out of 329 CIS students with baseline data. Each matched CIS student was matched with 19 comparison students, on average.

Table D.4 presents summary statistics showing how well CIS students were matched to comparison students on baseline characteristics. On average, comparison students were not significantly different from the CIS students on any baseline characteristics used in the analyses. Similarly, there were no statistically significant baseline differences between CIS and matched comparison students in the analytical samples used to estimate impacts for any outcome.

Impact model

To measure impacts using the matched sample, we estimated an OLS regression model that accounted for any small remaining differences between CIS and comparison students in their prior academic performance and characteristics:

(D1) $$y_i = \alpha + X_i\beta + \delta T_i + grade\ dummies + school\ dummies + \varepsilon_i$$

where y_i is the outcome of interest for student i; X_i is a vector of demographic controls, baseline test scores, and baseline test score imputation indicators for student i; T_i is a binary variable for treatment status, indicating whether student i received case management support from a CIS site coordinator; ε_i is a random error term that reflects the influence of unobserved factors on the outcome; and δ and β are parameters or vectors of parameters to be estimated, with δ representing the impact of the intervention of interest. Because each comparison student could be matched to multiple CIS students, we used a weighting scheme in which each CIS student had a weight of one, and each comparison student had a weight representing the fraction of the number of matching CIS students.

We assessed the impact of CIS case management on students' likelihood of being suspended after October 2017 or chronically absent in the 2017–2018 school year, as well as their performance on the Georgia Milestones and STAR exams from spring 2018. Compared to the analyses on the Georgia Milestones exam, the analyses of student performance on STAR ELA included 41 fewer students; the analyses of student performance on STAR math included 38 fewer students. No students were missing nonacademic outcome data (see Table D.2).

To interpret the estimated impacts on test scores, measured in standard deviations, we examined how the academic performance of the average student would change as a result of these impacts. We calculated the average *z*-scores of CIS students on the fall 2017 STAR

[51] The matching radii for CIS ranged from 0.1 to 0.35. We selected the radii for each grade to improve the quality of matches obtained.

assessment and added the impact estimates for the corresponding subject. We then used the standard normal distribution to determine the corresponding percentile ranks of those *z*-scores.

Table D.4. Baseline characteristics of matched CIS and comparison students

	CIS students	Matched comparison
2015–2016 CAAS reading *z*-score	-0.480 (0.775)	-0.470 (0.751)
2015–2016 CAAS math *z*-score	-0.419 (0.842)	-0.427 (0.834)
2016–2017 STAR ELA or SEL *z*-score	-0.641 (0.611)	-0.656 (0.619)
2016–2017 STAR math *z*-score	-0.538 (0.834)	-0.567 (0.833)
Fall 2017 STAR ELA or SEL *z*-score	-0.595 (0.609)	-0.612 (0.596)
Fall 2017 STAR math *z*-score	-0.450 (0.806)	-0.478 (0.804)
Black	0.973	0.973
Hispanic	0.027	0.025
Other race	0.038	0.031
Female	0.534	0.529
2016–2017 characteristics		
Homeless	0.038	0.041
Disabled	0.051	0.059
English language learner	0.021	0.019
Enrolled less than a full year	0.099	0.111
Behind grade level for his/her age	0.092	0.100
Attendance rate	0.948 (0.043)	0.946 (0.049)
Chronic absence	0.127	0.130
Ever suspended	0.116	0.123
Ever suspended in fall 2017	0.082	0.090
Suspended 2 or more days	0.058	0.067
Ever disciplined	0.209	0.209
Ever disciplined in fall 2017	0.116	0.122
Ever disciplined for a serious offense	0.110	0.110
Received academic support in 2017–2018	0.236	0.245
2015–2016 characteristics		
Homeless	0.086	0.069
Disabled	0.041	0.049
English language learner	0.027	0.024
Enrolled less than a full year	0.199	0.211
Behind grade level for their age	0.086	0.090
Attendance rate	0.944 (0.051)	0.942 (0.049)
Chronic absence	0.147	0.153
Ever suspended	0.075	0.065
Suspended 2 or more days	0.024	0.029
Ever disciplined	0.137	0.121
Ever disciplined for a serious offense	0.086	0.078
Number of students	**292**	**1,225**

Source: APS administrative data.

CAAS = Computer Adaptive Assessment System; ELA = English language arts; SEL = STAR early literacy.

Supplemental results on analyses presented in the main text

In addition to the main impact analyses summarized in the main text, we conducted exploratory analyses to assess whether the impacts differed for specific groups of students. Specifically, we tested whether the impacts of CIS case management varied for students who were served by CIS site coordinators with smaller caseloads, who thus could have dedicated more time to their students.[52] In addition, we tested whether the impacts differed for students who were "higher risk" in the prior school year due to having been either suspended or chronically absent in the 2016–2017 school year.

For each variable of interest, we estimated the following regression model, which adds interaction terms to the benchmark model in equation D1.

(D2) $$y_i = \alpha + X_i\beta + \delta T_i + \gamma T_i E_i + grade\ dummies + school\ dummies + \varepsilon_i$$

The coefficient γ represents how the impact differs for the exploratory variable of interest (E_i) (for example, whether the student was in a school with a smaller CIS caseload). Because students were not randomly assigned to schools, nor was their baseline discipline or attendance random, these analyses are exploratory and might reflect the influence of other related but unobserved factors.

The impact of CIS on students in schools with a smaller CIS caseload

Five of the targeted schools were classified as smaller caseload schools (BAMO, Boyd, Scott, Finch, and Kimberly Elementary Schools). We tested whether the impacts of CIS case management differed between CIS students who attended these five schools and those who attended schools with larger CIS caseloads, hypothesizing that site coordinators with smaller caseloads were able to give more attention to their students. The results of this analysis are presented in Table D.5. The results suggest that CIS case management had limited impacts regardless of the site coordinators' caseload size.

[52] We calculated the ratio of CIS students to CIS site coordinators in each school. If that ratio was smaller than 25 students per CIS site coordinator, we classified the school as a smaller caseload school. BAMO Elementary School had a full-time CIS site coordinator and Hollis Innovation Academy had two CIS site coordinators during the 2016–2017 school year. We divided their ratio by two before grouping the schools by caseload size. All other schools had one CIS site coordinator.

Table D.5. Impacts of CIS case management, by caseload size

	Academic outcomes (standard deviations)				Nonacademic outcomes (percentage points)	
	STAR ELA	STAR Math	GA Milestones ELA	GA Milestones Math	Likelihood of being suspended	Likelihood of being chronically absent
Impact of CIS case management services in larger caseload schools	-0.05 (0.04)	-0.04 (0.05)	-0.02 (0.04)	-0.03 (0.03)	0.01 (0.03)	-0.03 (0.03)
Impact of CIS case management services in smaller caseload schools	0.03 (0.04)	0.02 (0.05)	-0.06 (0.05)	-0.02 (0.04)	0.08 (0.05)	0.01 (0.03)
Difference in impacts between larger and smaller caseload schools	0.08 (0.05)	0.05 (0.07)	-0.04 (0.06)	0.01 (0.06)	0.07 (0.05)	0.04 (0.05)
Number of students	**1,220**	**1,215**	**1,456**	**1,457**	**1,517**	**1,517**

Source: APS administrative data.

Notes: This table displays impact estimates in z-scores (standard deviations) for the spring 2018 STAR assessment taken by students in all grades and the 2018 Georgia Milestones exams taken by students in grades 3–5, and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the likelihood that a student was ever suspended during the school year after October 2017. "Chronically absent" refers to the likelihood that a student had missed 10 percent or more of days enrolled. "Larger caseload schools" refers to schools where the ratio of students to CIS site coordinators was 25 or more students. "Smaller caseload schools" are schools where that ratio was smaller than 25. Standard errors are displayed in parentheses below each impact estimate. The sample size reflects the total number of CIS students and matched comparison students in each analysis. Differences in impacts may differ due to rounding.

**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

CIS = Community in Schools; ELA = English language arts.

The impact of CIS on high-risk students

Approximately 25 percent of all CIS students had been suspended or chronically absent in the 2016–2017 school year. As shown in Table D.6, we tested whether the impacts differed for these higher-risk students compared to other CIS students who had not been suspended or chronically absent in the previous school year to assess whether CIS case management services could benefit students with different levels of needs. There was little evidence that CIS case management was effective for either type of student.

Table D.6. Impacts of CIS case management on academic and nonacademic student outcomes, by high-risk status

	Academic outcomes (standard deviations)				Nonacademic outcomes (percentage points)	
	STAR ELA	STAR Math	GA Milestones ELA	GA Milestones Math	Likelihood of being suspended	Likelihood of being chronically absent
Impact of CIS case management on low-risk students	-0.01 (0.03)	-0.05 (0.04)	-0.02 (0.04)	-0.03 (0.03)	0.03 (0.02)	-0.04 (0.02)
Impact of CIS case management services on high-risk students	-0.06 (0.05)	0.07 (0.08)	-0.08 (0.07)	0.00 (0.06)	0.06 (0.06)	0.03 (0.06)
Difference in impacts between high- and low-risk students	-0.05 (0.06)	0.12 (0.09)	-0.06 (0.08)	0.04 (0.07)	0.03 (0.06)	0.07 (0.06)
Number of students	**1,220**	**1,215**	**1,456**	**1,457**	**1,517**	**1,517**

Source: APS administrative data.

Notes: This table displays impact estimates in z-scores (standard deviations) for the spring 2018 STAR assessment taken by students in all grades and the 2018 Georgia Milestones exams taken by students in grades 3–5, and in percentage point units for the suspended and chronically absent outcomes. "Suspended" refers to the likelihood that a student was ever suspended during the school year after October 2017. "Chronically absent" refers to the likelihood that a student had missed 10 percent or more of days enrolled. "Low-risk students" refers to students who had not been suspended nor chronically absent in the previous school year. "High-risk students" are those who had been suspended, chronically absent, or both in the previous school year. Standard errors are displayed in parentheses below each impact estimate. The sample size reflects the total number of CIS students and matched comparison students in each analysis. Differences in impacts may differ due to rounding.

**Impact is statistically significant at the 1 percent level.

*Impact is statistically significant at the 5 percent level.

CIS = Community in Schools; ELA = English language arts.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ SEATTLE, WA ■ TUCSON, AZ ■ WASHINGTON, DC ■ WOODLAWN, MD

MATHEMATICA
Policy Research