Building Assets and Reducing Risks (BARR) I3 Scale-Up Evaluation

Final Report

Johannes M. Bos, Cheryl Graczewski, Sonica Dhillon, Amelia Auchstetter, Julia Casasanto-Ferro, Sami Kitmitto

June 2022

American Institutes for Research® | AIR.ORG

AIR®
Advancing Evidence.
Improving Lives.

Building Assets and Reducing Risks (BARR) I3 Scale-Up Evaluation

Final Report

June 2022

Authors:

Johannes M. Bos, Cheryl Graczewski, Sonica Dhillon, Amelia Auchstetter, Julia Casasanto-Ferro, Sami Kitmitto

With *Ashley Sunde, So Jung Park, Briana Garcia, Connie Chandra*

American Institutes for Research®
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
202.403.5000 | **AIR.ORG**

Contents

Exhibits

Acknowledgments

A large-scale, complex evaluation like this one would not have been possible without the support and collaboration of many individuals and institutions. First, we want to thank the thousands of high school students who participated in the evaluation, responding to our surveys, participating in PSAT assessments, and allowing us to observe their classrooms and study their grades and attendance records. The support of these students and their parents is critical to this work in so many ways that it is difficult to overstate their contributions.

Second, we want to thank the teachers, principals, Building Assets and Reducing Risks (BARR) coordinators, and other school staff in the schools that participated in this study. These school staff and leaders allowed us to rearrange their ninth-grade master schedule, randomize their students, collect administrative data, and interview teachers and staff. They also facilitated communication with parents and other stakeholders in their districts and communities. Strong partnerships with schools are critical to the success of education evaluations, and these schools were exemplary partners throughout this work.

We also want to acknowledge the U.S. Department of Education's Investing in Innovation (i3) program for their generous financial support. This program, operated by the Office of Innovation and Improvement, was a major driver of innovation and rigorous effectiveness research in the U.S. education system. The i3 Scale Up grant awarded to the BARR Center enabled BARR to implement its program in the 66 study schools in this evaluation throughout the United States. The grant also paid for all our data collection, analysis, and reporting.

Throughout the design and execution of this evaluation, we received technical support from Abt Associates, whose representative, Marc Moss, was a strong and helpful technical reviewer of our work.

Within AIR, we want to acknowledge the support and guidance of Dean Gerdeman and Jordan Rickles, our quality assurance reviewers, and Andrew Wayne and Michael Vaden-Kiernan, our technical project reviewers. We also want to acknowledge our team members Emily Collins, Samantha LeVangie, George Atupem, Rouguia Barry, Femi Fadeyi, Cecilia Zhang, Keegan Giffels, Nick Kean, Feng Liu, Marina Castro, Evelyn Deng, Angela Su, Thomas Fink, Crystal Aguilera, Ethan Adelman-Sil, Alexa Cinque, Hana Gebremariam, David Blumenthal, Montrischa Essoka, Grace Fleming, Robin Bzura, Brenna O'Brien, and Trisha Borman.

We also want to thank the BARR Center and its staff and leadership, including Angela Jerabek, the director and originator of BARR; Justin Barbeau; Janice Eldridge; Maryann Corsello; Megan Reder-Schopp; Anu Sharma; Susan Savell; and Rob Metz. They were genuine partners throughout this evaluation of their program, recruiting 66 schools for the evaluation, and they provided us with invaluable program data and insights. Their support was critical to maintaining the schools' engagement, compliance with study protocols, and collaboration with data collection.

Executive Summary

Introduction

Building Assets, Reducing Risks (BARR) is a comprehensive, strength-based approach that uses eight components and focuses on the use of real-time data to build intentional staff-to-staff, staff-to-student, and student-to-student relationships in schools. This report focuses on an evaluation of BARR in ninth grade, where BARR aims to facilitate the challenging transition from middle to high school. The developer designed BARR around the following eight interlocking components:

1. a focus on the whole student (i.e., each student's academic, emotional, social, and physical needs)
2. professional development for school staff (i.e., trainings and coaching support)
3. BARR's social-emotional curriculum (i.e., "I-Time") to foster teacher-to-student and student-to-student relationships and help students learn and practice life skills
4. cohorts of students and staff to help educators cultivate connections with students and each other
5. regular block/team meetings of the cohort teacher teams to collaboratively identify struggling students and interventions, as well as students who should be accelerated
6. risk review meetings with school leadership, support staff, and community resources to address the needs of students who need more intensive support
7. partnering with families in student success
8. engaging school administrators to achieve specific, measurable goals and to network with other administrators in the BARR community for ideas

This scale-up evaluation is the third in a series of randomized controlled trials funded by the U.S. Department of Education's Investing in Innovation (i3) program. It follows a 2010 development grant (Corsello & Sharma, 2015) and a 2013 validation grant (Bos et al., 2019; Borman et al., 2021). This evaluation documents the scale-up of the BARR model during the grant period (2017–2021) and the implementation and impacts of the program in its first year of implementation in 66 schools across the United States.

The 66 schools included in the impact evaluation were distributed across three cohorts (2017–18, 2018–19, and 2019–20 school years) and included 21,529 students and 524 teachers. The schools were recruited from 12 states and the District of Columbia, with a focus on relatively low-performing schools and districts. Of the ninth-grade students in this evaluation, 64 percent

were eligible to receive free or reduced-price lunch and 15 percent were English learners. Approximately one third of students were white, and two thirds were students of color. Of the 66 schools, 12 were rural.

For the evaluation, schools in each cohort were randomly selected to receive the BARR model immediately (the treatment group) or receive BARR after one year (a "wait-list" control group) to provide a reference for measuring the impacts of the BARR model.

Some Study Schools Experienced Implementation Challenges

Schools participating in this evaluation and assigned to receive the BARR model (i.e., the treatment group) experienced challenges implementing BARR with fidelity during the first year more so than schools implementing the model in the two prior evaluations. This was the case especially in Cohort 1 and Cohort 3 but for very different reasons. In Cohort 1, three of the 15 schools randomized to BARR decided during the summer before implementation that they were unable to implement BARR that year.[1] Similarly, one of the 11 treatment schools in Cohort 2 also decided not to implement BARR. In Cohort 3, implementation of BARR was off to a good start, only to be severely compromised by the COVID-19 pandemic, which closed all school campuses in the 11 BARR schools in March 2020. None of these campuses reopened during the 2019–20 school year, and all teaching happened virtually after the campuses closed. All BARR activities continued and were modified to accommodate virtual delivery and to meet the new needs of teachers and school administrators during this challenging time. After excluding schools randomized into BARR but unable to implement in the study year, 67 percent of schools in Cohort 1, 50 percent of schools in Cohort 2, and 50 percent of schools in Cohort 3 met BARR Center's criteria for implementation of the BARR model with fidelity.[2]

BARR Improved Teacher Experiences and Attitudes

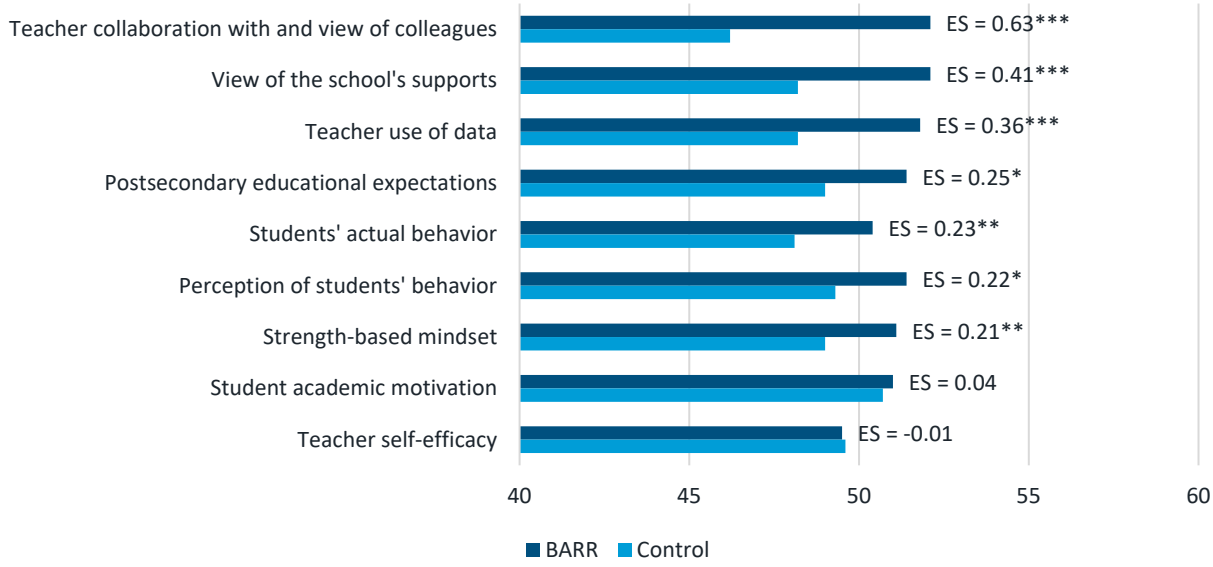
Teachers are a major lever in BARR's theory of action. This study is the first BARR evaluation in which teachers were randomized (with their schools). The BARR model changes teachers' work environment, and BARR provides them with professional development and coaching on the BARR components introduced previously. It also sets expectations for teacher-to-teacher collaboration and for developing a more holistic view of their students' strengths and experiences. Because of this, we expected that BARR would impact teacher experiences and attitudes.

[1] These schools became "crossover" schools for the purpose of the evaluation. That is, they remained classified as part of the BARR treatment group for analysis even though they were not able to implement and benefit from BARR during the study year but were ready to implement BARR the following year.

[2] BARR Center plans to re-examine the fidelity rubric and its scoring based on the results of this evaluation.

To capture these impacts, the evaluation assessed changes in teacher experiences along nine constructs (using a teacher survey) and found significant positive program effects for seven of these constructs (Exhibit 1). The most substantial effects were on "teacher collaboration with and view of colleagues," "teacher use of data," and "view of school supports." The corresponding effect sizes (ES[3]) were moderate to large (ES ranging from 0.36 to 0.63).

Exhibit 1. Impacts of BARR on Teacher Experiences and Attitudes



Note: All outcomes are survey scales with a mean of 50. "ES" is the effect size associated with the difference between the treatment and control groups. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level.

BARR Improved Credit Attainment and Reduced Course Failure

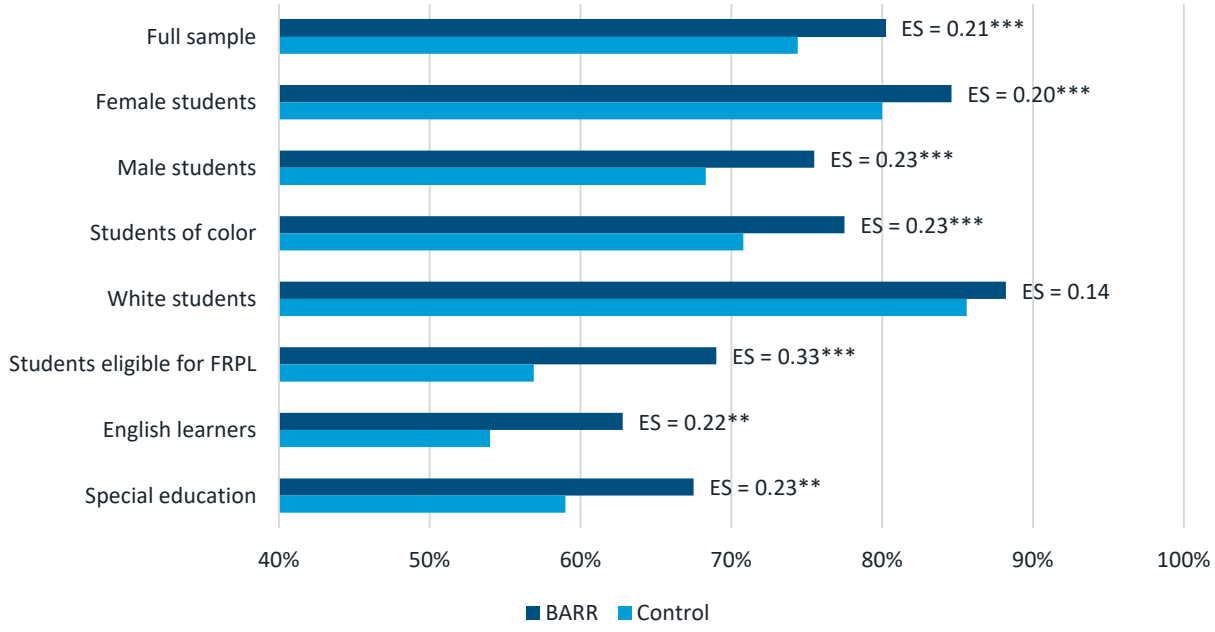
Reducing course failure in ninth grade is a major focus for BARR. The transition from middle to high school comes with new academic challenges that cause many students to begin experiencing course failures. Unless prevented or addressed, those failures have potentially serious consequences for students throughout their academic careers. By helping schools, teachers, students, and parents to stay on top of student performance during the ninth-grade year, BARR seeks to prevent students from failing courses and missing credits. This can also benefit their GPA.

As shown in Exhibit 2, across the three cohorts combined, BARR had substantial and statistically significant impacts on the proportion of students who passed all their core courses, an increase from 74 to 80 percent, for an effect size of 0.21. BARR's effects on credit attainment and course

[3] To calculate an effect size, we divide the impact by the pooled standard deviation of the outcome variable in the treatment and control groups. Effect sizes allow us to easily compare the size of estimated program effects across outcomes and studies.

failure were strongest for male students, students of color, students eligible for free or reduced-price lunch, English learners, and students with disabilities. These groups were more likely to experience course failure, as evident in the control group. BARR thus reduced existing gaps in these outcomes between different demographic groups of ninth-grade students.

Exhibit 2. Impact of BARR on Percentage of Students Passing All Core Courses, Full Sample and Student Groups



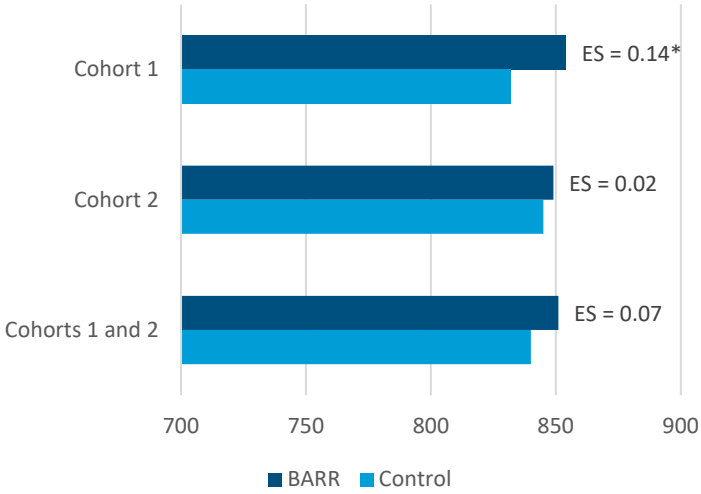
Note: "ES" is the effect size of the difference between the treatment and control groups. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level.

BARR's impact on students' GPAs (not shown) was more modest than its impact on credit attainment and course failure (an increase from 2.5 to 2.6, for an effect size of 0.13). BARR's impacts on GPA were stronger for students of color, students eligible for free or reduced-price lunch, and English learners.

BARR Improved Academic Achievement in Cohort 1

We used the Preliminary Scholastic Aptitude Test (PSAT), administered in the fall of tenth grade, as an independent measure of student achievement in the evaluation. Due to the COVID-19 pandemic, we could not consistently administer the PSAT test in fall 2020. Therefore, this outcome was available for impact analysis only for Cohort 1 and Cohort 2. Exhibit 3 shows that BARR had a positive impact on PSAT scores in Cohort 1 (an average score of 854 in BARR schools compared to 832 in control schools, for an effect size of 0.14). There was no such impact in Cohort 2, and the combined impact for Cohort 1 and Cohort 2 (an estimated difference of 11 points—840 in control schools and 851 in BARR schools) was not statistically significant.

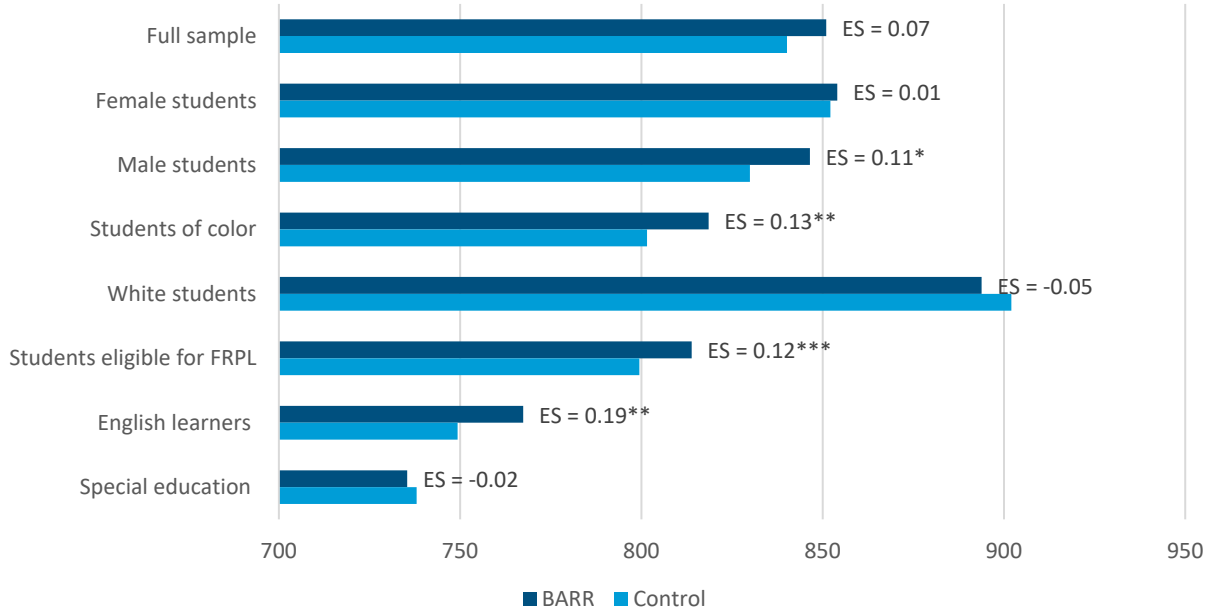
Exhibit 3. Impacts of BARR on PSAT Scores, by Cohort



Note. "ES" is the effect size of the difference between the treatment and control groups. * = statistically significant at the $p < .05$ level.

Analyses of PSAT scores for different groups of students (Exhibit 4) showed a significant positive impact on the PSAT for male students, students of color, students eligible for free or reduced-price lunch, and English learners. After adjusting for the fact that some schools did not implement BARR in the study year or did not offer it to all their students, we found positive impacts on the PSAT scores of male students, students of color, students eligible for free or reduced-price lunch, and English learners (ES = 0.13, 0.16, 0.14, and 0.21, respectively; not shown in exhibit).

Exhibit 4. Impacts of BARR on PSAT Scores, by Student Group

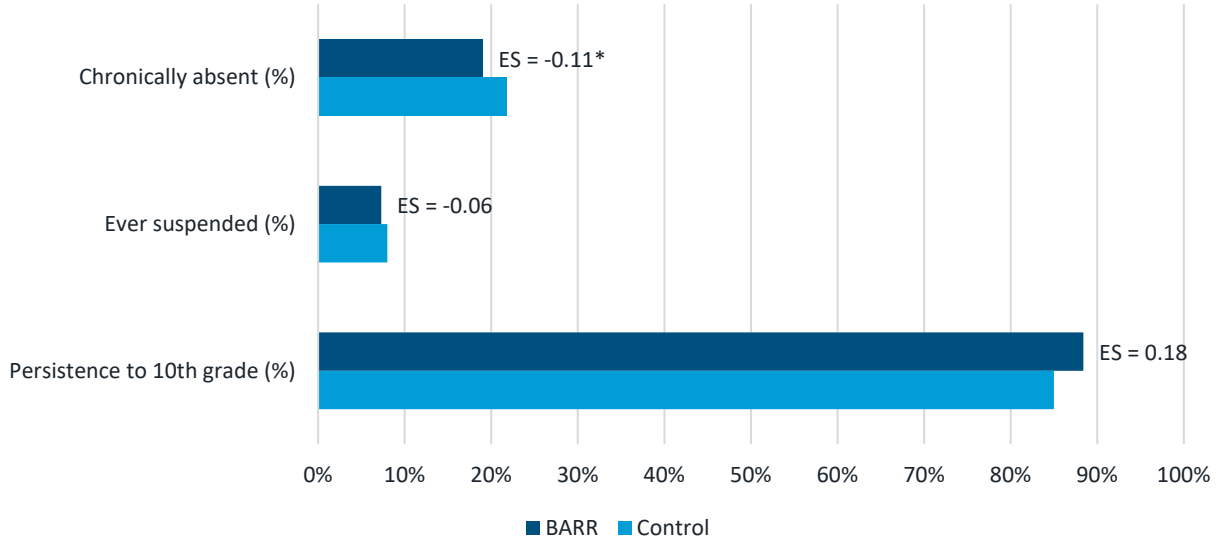


Note. "ES" is the effect size of the difference between the treatment and control groups. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level.

BARR Reduced Chronic Absenteeism

Examining behavioral outcomes as recorded in school administrative data, we found that BARR significantly reduced chronic absenteeism (students being absent more than 10 percent of the time—Exhibit 5). Assignment to BARR caused a negative (favorable) impact on the percentage of students who were chronically absent. The rate of chronic absenteeism was 19 percent in BARR schools, compared to 22 percent in control schools (ES = –0.11).

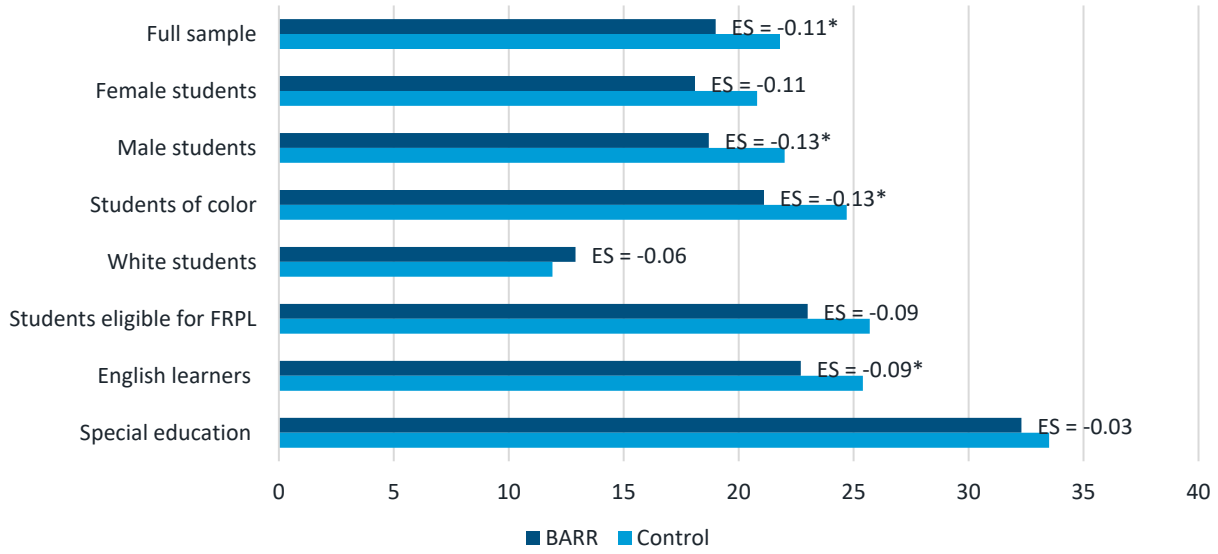
Exhibit 5. BARR Impact on Chronic Absenteeism, Suspension, and Persistence to 10th Grade



Note. "ES" is the effect size of the difference between the treatment and control groups. * = statistically significant at the $p < .05$ level.

Examining the impact on chronic absenteeism by student group, we found that impacts on chronic absenteeism were strongest for male students and students of color as shown in Exhibit 6 (ES = –0.13 for both groups).

Exhibit 6. Impact of BARR on Chronic Absenteeism, by Student Group



Note. "ES" is the effect size of the difference between the treatment and control groups. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level.

BARR Impacts on Suspensions and Persistence to 10th Grade Were Mixed

As shown in Exhibit 5, we did not find statistically significant impacts on suspension rates when examining these outcomes across all three cohorts or on persistence to 10th grade in the same school when examining these outcomes across two cohorts. We did find a statistically significant reduction in suspensions in Cohort 1 (from 11.6 to 7.6 percent of students—not shown) but no comparable impacts in the other two cohorts.

Scale-Up of the BARR Model Was Successful

During the grant period, BARR significantly increased the number of schools it served, from 26 in the 2015-16 school year (the last year before the scale-up grant) to 224 in the 2021-22 school year, the last year of this grant. To do so, the BARR developer built a sustainable organizational infrastructure, including *BARR Center*, a 501C3 not-for-profit corporation from which to manage the work. With this center as its base, BARR and its partner Hazelden Betty Ford Foundation assembled a national network of regional coaches and mentors, who have assumed the responsibility for providing day-to-day support to BARR schools and their districts. As a result, the original BARR developer and her team at BARR Center are no longer solely responsible for most of the day-to-day interactions and support activities with BARR schools but rather provide oversight, which has proven important to scaling the model nationwide.

BARR Center demonstrated progress in scaling six strategies that they considered important to grow and sustain the model over time. During the grant period, BARR Center was particularly successful in strengthening BARR Center infrastructure (i.e., sufficient staffing, sustainable budget), determining and sharing the cost of services, and scaling school-level and coach training and supports. These strategies were under direct control of BARR Center, and their successful scaling was facilitated by attentive and responsive BARR Center staff and coaches and a focus on building infrastructure. Encouraging schools to build local awareness of the BARR model and to continue implementing the model over time were harder to scale. These two strategies were primarily driven by school administrator decisions and not within the direct control of BARR Center.

BARR Center is working to build additional infrastructure and supports to address remaining scale-up challenges. For example, BARR Center implemented a mentorship model to train new coaches and a coach mastery rubric that examines coaches' proficiency in essential knowledge areas to ensure newly trained BARR coaches have the tools and capacity needed to support the diverse needs of schools working toward full implementation of the BARR model. BARR Center also plans to launch a school accreditation process that formally certifies schools that continue to implement the BARR model over multiple years.

Conclusion

The findings of this scale-up evaluation are consistent with those we found in our 2019 i3 validation study of BARR and the ones that Corsello and Sharma found in their 2015 evaluation of BARR's first i3 development grant. All three studies found substantial positive impacts on credit attainment and course failure and smaller effects on academic achievement and grades. In this evaluation, we also found a favorable effect on chronic absenteeism and substantial favorable effects on a range of teacher experiences and attitudes.[4]

In this evaluation, more so than in the two previous studies, there was considerable variation in impacts across cohorts and student groups. Partially, this may reflect underlying differences in the composition of the three cohorts of schools and partially it reflects differences in study school implementation challenges. Cohort 3 was significantly impacted by the COVID-19 pandemic.

The findings from this evaluation continue to support the conclusion we made at the end of the validation study, which is that BARR is an effective model for schools aiming to improve students' transitions into ninth grade, reduce course failure, and narrow gaps in student academic outcomes between different demographic groups of students, while improving teacher collaboration and view of colleagues, use of data, and view of school supports.

[4] The other two BARR i3 evaluations also found that BARR teachers were more positive about their experiences than control teachers, but in those two studies, teachers were not randomly assigned to BARR. This study is the first BARR evaluation in which teachers were randomized (with their schools).

Chapter 1. Introducing BARR and the Study Design

Overview of the BARR Model and Evaluation Design

Building Assets, Reducing Risks (BARR) is a comprehensive, strength-based approach that uses eight interlocking components and focuses on the use of real-time data, building intentional staff-to-staff, staff-to-student, and student-to-student relationships. Based on previous research with the model (e.g., Borman et al., 2021; Bos et al., 2019; Corsello & Sharma, 2015), schools that implement the program with fidelity should expect to see notable changes in teachers' perceptions of self, students, and the school environment as well as students' school experiences, behaviors, and, over time, academic outcomes. After completing successful development and validation grants, the U.S. Department of Education's Investing in Innovation (i3) program provided BARR with a scale-up grant to bring the BARR model to more high schools around the country. As part of this grant, the American Institutes for Research (AIR) conducted an independent evaluation of the impact of the BARR model on Grade 9 student outcomes in 66 high schools. Within each of these schools, the evaluation focused on the first year of implementation, measuring implementation fidelity and teacher- and student-level outcomes at the end of Year 1 of this three-year program. The schools participated in the study in three distinct one-year cohorts (22 in 2017–18, 22 in 2018–19, and 22 in 2019–20).

The impact evaluation uses a school-level randomized controlled trial (RCT) to estimate the impact of the BARR model. AIR randomly assigned high schools participating in the evaluation to the BARR model or to a "business as usual" (BAU) control group. This evaluation measured the impact of the resulting treatment contrast on a range of academic and nonacademic outcomes after the first year of implementation.

In addition to the impact evaluation, the i3 grant also includes a focus on examining the conditions that grantees must address to grow their program over time and sustain long-term positive program outcomes for schools that participate in the program. As part of our evaluation, we worked with BARR Center to document the strategies that BARR Center identified as important to grow and sustain the program as well as BARR Center activities related to these strategies.

This report first provides a brief description of the BARR model and the study design, including the research questions, characteristics of the schools and students participating in the study, and the measures and analytic approach we used to address the research questions. Subsequent sections describe the implementation of BARR by the study schools, impacts on teacher experiences and attitudes and students' academic and behavioral outcomes, and

efforts by BARR to grow and sustain the program. We conclude with limitations and a summary of findings.

What Problem Does BARR Address in this Scale-up Grant?

The transition from eighth grade to ninth grade is a critical point for students that can set them on a path either toward academic success or failure. Upon entering high school, students encounter a wealth of new people, content, structures, and expectations. With this abrupt change, it is common for students to experience increased stress and behavior problems alongside declines in grades, attendance, interest in school, and perceptions of academic competence and self-esteem (Alvidrez & Weinstein, 1993; Benner & Graham, 2009; Reyes et al., 2000; Uvaas & McKevitt, 2013). These problems not only affect students during ninth grade, but also can impact their long-term high school success. Ninth-grade attendance and course performance, for instance, have been shown to be highly predictive of high school graduation or dropout (Allensworth & Easton, 2005; Roderick et al., 2014; Neild et al., 2008).

Support for the BARR Model

Many programs designed to address the challenges facing students entering ninth grade have attempted to improve the student experience by changing the organization of the ninth-grade year. Examples of such structural changes include maintaining a "blocked" teaming structure like the one that many students experience during their middle school years by organizing ninth-grade students into teams with common groups of students and teachers. Research shows that students enrolled in these blocked settings have better course performance than students enrolled at traditional high schools (Ellerbrock & Kiefer, 2013; Eccles & Roeser, 2011; Styron & Peasant, 2010; Cook et al., 2008; Neild, 2009). Another promising structural approach to helping ninth-grade students transition into high school is to improve the flow of information among educational staff about how these students are doing across all their classes. This includes the development and use of "early warning" data tools (e.g., Faria et al., 2017; Corrin et al., 2016; Allensworth, 2013). These tools take advantage of the growing adoption of real-time integrated data systems to identify areas of concern and refer students to appropriate interventions.

The Importance of Relationships. Underlying these and other interventions targeting ninth-grade transitions specifically and secondary school reform more broadly is the belief that relationships are critical to student success. Research has shown that students' academic and nonacademic experiences and outcomes are influenced by three major relationship sources: parents/guardians, teachers, and peers. Better academic functioning has been linked to parents' positive expectations and academic goals for their child, the quality and consistency of parental feedback on the child's behavior and performance, the educational values and standards that parents hold for their child, and direct parental engagement in the child's

education (Martin & Downson, 2009; Mansour & Martin, 2009; Bempechat & Shernoff, 2012; Pomerantz & Moorman, 2010). Similarly, research has documented strong relationships between the quality of student-teacher relationships and students' academic and nonacademic development and motivation (Scales et al., 2020; Martin & Collie, 2019; Juvonen et al., 2012; Kontos & Wilcox-Herzog, 1997; Teven & McCroskey, 1997; Connell & Wellborn, 1991). Last, research has associated positive peer (student-to-student) interactions with higher levels of motivation, greater engagement in school, and better academic performance (Benner et al., 2017; Liem & Martin, 2011; Furrer & Skinner, 2003). Research also suggests that teachers play an important role in facilitating the development of relationships among their students and that strong peer relationships can mediate the effects of teacher emotional support and positive teacher-student relationships on student academic outcomes (Ruzek et al., 2016).

The Importance of a Supportive Work Environment. Researchers have used the interconnectedness of the social, academic, and affective dimensions of the student to argue that teachers and educational programs must focus on 'the whole self' within the academic context (Osher et al., 2018; Sibley et al., 2017; Martin & Dowson, 2009). To build positive relationships with students, teachers therefore must learn more about their students' personal experiences and needs and their performance in other subjects and school activities. As they do so, teachers may feel overwhelmed or powerless if they feel that they must address their students' needs on their own. This may cause teachers to experience exhaustion, reduced empathy, a lack of feelings of personal accomplishment, and depression (Halbesleben, 2006; Jennings & Greenberg, 2009; Mahan et al., 2010). To address these challenges and build and maintain strong relationships with their students, it is therefore important that teachers have positive and supportive working relationships with other teachers, administrators, and other school staff. Collaborative working environments, such as professional learning communities (PLCs), tend to foster teacher collaboration and a shared focus on student learning, reinforce teacher authority, and promote continuous professional learning (Vescio et al., 2008). Such environments serve as a mechanism for building supportive relationships among teachers, which may help teachers persevere when faced with feelings of isolation, exhaustion, and job dissatisfaction (Beltman et al., 2011). Furthermore, research shows that teachers who work in more supportive and collaborative environments are more likely to improve their effectiveness over time than those who are in less supportive environments (Johnson et al., 2012). Indeed, high-quality and sustained collaboration among teachers is associated with greater student achievement (Ronfeldt et al., 2015).

Development and Evaluation of the BARR Model

The BARR model was initially implemented in St. Louis Park High School, Minnesota, and has been in operation there since the 1998–99 school year. It was initially funded through a Minnesota

Department of Human Services (Center for Substance Abuse Prevention) State Incentive Grant. With subsequent funding from a 2010 i3 development grant, BARR was then implemented in two schools in Maine and one school in California. This implementation was rigorously evaluated with an RCT in the California school, which found statistically significant positive effects on reading achievement, mathematics achievement, the number of core credits earned, grade point average, and the course failure rate (Corsello & Sharma, 2015).

Based on these results, BARR was awarded an i3 validation grant in 2013 to support a rigorous study of the BARR model in 11 schools in Maine, California, Minnesota, Kentucky, and Texas. This study included 4,168 Grade 9 students, the majority of whom were eligible for free or reduced-price lunch (75 percent) and were white (75 percent). For the total student sample, the validation study found that BARR increased the percentage of core credits earned (effect size [ES] = 0.19) and the average Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) mathematics scores (ES = 0.11) (Borman et al., 2021).

Following promising results from the first two cohorts of schools in the validation study, in 2017, BARR applied for and was awarded the i3 scale-up grant that is the subject of this evaluation in 2017. As part of the scale-up process, the original BARR developer and her team at the BARR Center are no longer solely responsible for most of the day-to-day interactions and support activities with BARR schools but rather provide oversight, which has proven critical to scaling the model nationwide.

BARR has grown substantially over the years and, as of this writing, BARR Center has implemented the model in more than 180 schools.

Components of the BARR Model

1. Focus on the whole student,
2. Professional development,
3. "I-Time" social-emotional lessons,
4. Restructured course schedule,
5. Regular block/team meetings among core subject teachers,
6. Risk review meetings with school leadership and support staff,
7. Engaging families, and
8. Engaging school administrators.

Components of the BARR Model

BARR is designed around eight interlocking components to build intentional staff-to-staff, staff-to-student, and student-to-student relationships. These components include (1) a focus on the whole student (i.e., including all the assets, resources, challenges, and risks they bring to school); (2) professional development for teachers, counselors, and administrators; (3) a series of "I-Time" social-emotional lessons to improve classroom relationships and foster a supportive climate of learning; (4) a restructured course schedule to create groups of students with common teachers; (5) regular block/team meetings among core subject teachers to collaboratively problem-solve and encourage student progress, focusing on all students; (6) risk review meetings with school leadership, support staff, and community resources to address the needs of students

with more persistent or serious challenges; (7) partnering with families to foster student success; and (8) engaging school administrators to support teachers and take an active role in student achievement.

To implement the eight components, the BARR model provides professional development for BARR teachers, administrators, counselors, and the BARR coordinator in each school prior to the school year. The BARR coordinator, who is a school staff member, receives a BARR coach who provides professional development and coaching for approximately 200 hours over the course of the school year. In addition, there are monthly meetings of BARR teachers, quarterly PLC calls, and an annual conference for all BARR educators. Partnering with families is another component of implementing the model that begins with a summer orientation followed by an invitation for all parents or guardians of students to join a parent advisory committee. In addition to family engagement, ensuring district-level and building-level administrative support for implementation of the model is an important component of the BARR model.

The BARR coach works with the school to create the structural and organizational conditions necessary to fully integrate student supports into a school's existing model for addressing academic and nonacademic barriers to learning. Implementation of the BARR model within a school involves restructuring Grade 9 so that groups of students share the same group of core subject teachers. This restructuring includes finding a common time for the group of core subject teachers to meet to discuss all students' strengths and challenges. During these block/team meetings, teachers share their individual experiences with students and collectively review real-time student data to identify interventions that may be helpful. If students appear to be in crisis or do not respond to block/team meeting interventions, they are referred to a "risk review" process. Risk review meetings convene both academic and nonacademic staff (e.g., counselors, assistant principals, school resource officers) and community resources to identify more targeted interventions. Both academic and nonacademic staff are included so that there is a shared understanding of the assets the students bring to school, the academic and nonacademic risks they face, and the interventions that are underway to support them. In risk review meetings, school staff and community liaisons engage in a collaborative assessment of the progress and challenges of high-risk students and problem-solve to address these challenges.

In the BARR model, teachers also deliver a weekly "I-Time" lesson (on a rotating basis) to all students. These lessons focus on students' social and emotional development and encourage students and teachers to build relationships with each other by sharing insights and experiences. The model also emphasizes support for the whole student, which means that teachers work collaboratively through the block meeting process and the delivery of I-Time to develop an understanding of working with the whole student. Consequently, they expand their

perception of their role as teachers, from merely providing academic instruction to supporting all facets of students and working collaboratively with their peers. We discuss each of these eight components in more detail in Chapter 2.

Together, these eight components intend to improve the high school experience for students (e.g., feeling more connected to school, cultivating better relationships with teachers, receiving coordinated support) and for teachers (e.g., developing better relationships with colleagues, working collaboratively, feeling empowered to support students). BARR expects these improved experiences to translate into better short-term outcomes for students (e.g., earning more course credits toward graduation, attaining better test scores, being more engaged in learning) and eventually to result in long-term benefits for students (e.g., increased graduation rates, higher college acceptance rates). A logic model tying these different outcomes together is provided in Appendix A2.

Study Design

To examine the impact of BARR, we implemented an RCT that used school-level random assignment of 66 schools in 12 states and the District of Columbia, including 512 teachers and 21,529 students. The schools agreed to be randomly assigned to implement BARR or to a business-as-usual control group. We also designed this study to examine (1) how BARR scaled the program during the grant award period, and (2) how well schools assigned to BARR implemented program components during the study year. Student outcomes include credit attainment, grade point average (GPA), academic achievement, attendance, suspensions, and persistence to 10th grade within the same school. We also measured students' attitudes and experiences in ninth grade and a range of teacher outcomes, including teachers' perception of their performance, their colleagues, and their students.[5]

Research Questions

Exhibit 1.1 shows the evaluation research questions. The first set of research questions focuses on the scaling of the BARR model during the grant award period and the implementation of BARR in treatment schools.

[5] Due to survey nonresponse and related methodological challenges, the student survey data did not meet our data quality standards and were not included in the impact analyses presented in this report. Please see Appendix F for details.

Exhibit 1.1. Research Questions

Scaling and Implementation Research Questions
What are the facilitators and barriers to successful scale-up of the BARR model, and how are they addressed/overcome?
To what extent are the BARR scale-up strategies implemented as intended?
What are the facilitators and barriers encountered to successful school implementation of the BARR model in the study treatment schools?
To what extent is the BARR model implemented as intended in the study treatment schools?
Educational Attainment and Student Achievement Impact Research Questions[6]
What is the impact on **ninth-grade students' educational attainment** for those students attending schools that were assigned to the BARR model compared to students in the business-as-usual (BAU) condition?
What is the impact on **10th-grade students' reading performance** (after one year of exposure in ninth grade) for students in schools assigned to the BARR model compared to students in schools assigned to the BAU condition?
What is the impact on **10th-grade students' mathematics performance** (after one year of exposure in ninth grade) for students in schools assigned to the BARR model compared to students in schools assigned to the BAU condition?
What are the impacts of the BARR model on **ninth-grade students passing all core courses** for students attending schools that were assigned to the BARR model compared to students in the BAU condition?
What are the impacts of the BARR model on **ninth-grade students' grade point average** at the end of ninth grade for students attending schools that were assigned to the BARR model compared to students in the BAU condition?
What are the impacts of the BARR model on each outcome for students attending schools that were assigned to the BARR model compared to students in the BAU condition across different groups of students?
Teacher Experience Research Question
What are the impacts of the BARR model on **ninth-grade teachers' perceptions of self, their students, and their school** for teachers working in schools assigned to the BARR model compared to teachers in the BAU condition?
Student Behavior Research Questions
What are the impacts of the BARR model on **ninth-grade students' chronic absenteeism** for students attending schools that were assigned to the BARR model compared to students in the BAU condition?
What are the impacts of the BARR model on **ninth-grade students' suspensions** for students attending schools that were assigned to the BARR model compared to students in the BAU condition?
What are the impacts of the BARR model on **ninth-grade students' persistence to 10th grade** for students attending schools that were assigned to the BARR model compared to students in the BAU condition?

[6] The first three research questions under Educational Attainment and Student Achievement are the evaluation's confirmatory research questions.

The second set of research questions addresses the impact of the BARR model on student educational attainment and student academic achievement.[7] Additional questions consider several other student academic outcomes (i.e., passing all core courses, GPA). We also explored the extent to which impacts varied by school-reported student background characteristics, including gender, whether students are students of color, English learner (EL) status, special education status, and whether students qualified for free or reduced-price lunch. The third set of research questions explores the extent to which assignment to the BARR model impacted teachers' perceptions of self, their students, and their school. The last set of research questions explores the BARR model's impact on student behaviors (i.e., absenteeism, suspensions, persistence to 10th grade).

Sample

This study included three cohorts of schools from 12 states and the District of Columbia. BARR Center recruited 22 schools for each cohort to participate in the study.[8] Consistent with initial guidance from the i3 grant program, BARR Center's recruitment efforts for Cohort 1 (implementing in 2017–18) focused on schools that were considered among the lowest performing in their respective states. During the 2017–18 school year, BARR Center noticed that some of the recruited Cohort 1 schools had limited infrastructure to implement the program.[9] They subsequently established minimum infrastructure requirements for schools to be selected for cohorts 2 (implementing in 2018–19) and 3 (implementing in 2019–20). Consequently, although cohorts 2 and 3 included similarly high-need schools as Cohort 1, the schools included in these cohorts had to meet stricter infrastructure requirements and were more closely scrutinized for their readiness to implement BARR.

Exhibit 1.2 compares the school and student background characteristics of schools and the student background characteristics of the students in the sample in the three cohorts. In total, the study includes 66 schools (37 treatment and 29 control) across three cohorts with all students entering ninth grade in the sample schools. Exhibit 1.2 lists the number of schools and students assigned to treatment and control groups by cohort. The combined sample across the three cohorts of schools included 21,529 ninth-grade students: 12,122 BARR students and 9,407 control

[7] The first three research questions in the second panel in Exhibit 1.1 are the confirmatory research questions that cover what BARR Center considered to be the most essential summative outcomes of the BARR model. That is, a positive impact on one or more of these outcomes should be considered evidence of the intervention's overall effectiveness. The number of confirmatory impact questions is limited in studies like these to reduce the likelihood of finding a statistically significant effect by chance. These three confirmatory questions were preselected before any outcome data were collected. Please see the What Works Clearinghouse Standards Handbook (Version 4.0) for additional information about the proper interpretation of confirmatory and exploratory impact estimates and related concerns about the statistical implications of multiple comparisons in impact evaluations (What Works Clearinghouse, 2020).

[8] The 66 schools recruited for the study were from the following 12 states and the District of Columbia: Arizona, California, Illinois, Maine, Maryland, Massachusetts, Minnesota, New York, Tennessee, Texas, Wisconsin, and West Virginia.

[9] Examples of infrastructure constraints included maintenance of the physical plant as well as limited access to technology for collecting and sharing data.

students. Exhibit 1.2 also provides the background characteristics of the teachers who completed the survey in the three cohorts. The combined teacher sample across the three cohorts of schools included 524 ninth-grade teachers: 286 BARR teachers and 238 control teachers. Additional details about school recruitment and selection, and randomization are provided in Appendix A.

Exhibit 1.2. Characteristics of Study Schools, Students, and Teachers in Cohorts 1, 2, and 3

	Cohort 1	Cohort 2	Cohort 3
School assignment at randomization	Count	Count	Count
Treatment schools	15	11	11
Control schools	7	11	11
School characteristics	Average	Average	Average
Percent students of color	45.8	73.9	62.6
Percent free or reduced-price lunch	62.9	68.1	50.0
Percent English learners	6.6	20.2	13.9
Percent special education	16.0	16.6	14.9
Total enrollment	1,135	1,266	1,321
Graduation rate	82.6	82.4	85.0
Student assignment at randomization	Count	Count	Count
Treatment students	4,246	3,610	4,266
Control students	2,059	4,042	3,306
Student characteristics	Percent	Percent	Percent
Students of color	58.1	79.0	63.7
Free or reduced-price lunch	59.1	71.7	59.3
English learners	7.6	25.6	12.4
Special education	16.2	15.5	14.2
Teacher assignment at randomization	Count	Count	Count
Treatment teachers	83	107	96
Control teachers	67	104	67
Teacher characteristics	Percent	Percent	Percent
Female	64.0	52.1	57.7
Teachers of color	22.7	24.2	19.0
Advanced degree	57.3	64.5	70.6
1 year experience	6.7	8.5	0.0
2-5 years' experience	29.3	27.5	24.5
6-10 years' experience	16.0	15.6	14.7
11+ years' experience	43.3	42.2	54.6

Source**:** Common Core of Data 2017–18, 2018–19, and 2019–20 used for school percent students of color, percent free-and-reduced lunch program, and total enrollment. State education agency websites used for percent EL and special education. US Department of Education's EDFacts initiative used for graduation rates. AIR administrative data for student counts and characteristics. AIR survey data for teacher counts and characteristics.

BARR Model Implementation Measurement

We observed in-school BARR meetings (i.e., I-Time lessons, block meetings, and risk review meetings) and collected interview data from school staff (BARR coordinators, teachers, and administrators) to understand the facilitators and barriers to program implementation encountered by treatment schools and to assess the extent to which the BARR model was implemented as intended by treatment schools. We coded these interview data to identify the facilitators of and barriers to implementing each component of BARR and the whole program. To assess the extent to which the BARR model was implemented as intended, we worked with BARR Center to establish thresholds for each of the evaluated implementation components, defining the level at which BARR Center considered the BARR component to be adequately implemented at a school. We used a structured interview process to ask BARR coordinators to rate their school on relevant indicators for each of the eight components of the BARR model. In addition, whenever possible, two evaluators directly observed and rated block/team meetings, I-Time lessons, and risk review meetings using the fidelity rubrics developed for the 2010 i3 development project (Corsello & Sharma, 2015) and adapted for this study.[10] These data were combined to calculate a fidelity score for each component. To meet adequate implementation, we established that schools needed to meet or exceed thresholds for at least four of five priority components and two of three of the remaining components. Additional details about the BARR model implementation data sources and measurement are provided in Chapter 2 and Appendix C.

Teacher and Student Outcome Measurement

We used three data sources to obtain the outcome measures that we used in the impact analysis: (1) a teacher survey to measure teachers' perception of self, students, and schools; (2) school administrative data to measure students' academic outcomes, attendance, suspensions, and persistence into 10th grade; and (3) a study-sponsored administration of the Preliminary SAT/National Merit Scholarship Qualifying Test (PSAT) to measure academic achievement. We administered the teacher surveys during the spring semester to all ninth-grade teachers.[11] Teacher surveys provided measures to evaluate the impact of BARR on teachers' perception of self, students, and schools. We requested and received administrative data, including demographic background data about the students from study schools, districts, and, in some cases, state education agencies. To supplement the academic measures in the administrative data, we asked schools to administer the PSAT during the fall semester of the students' 10th-

[10] We did not conduct observations at Cohort 3 schools due to COVID-19 restrictions. In cohorts 1 and 2, we occasionally experienced scheduling conflicts that limited our ability to have two evaluators in each observation.

[11] Surveys were administered in-person for cohorts 1 and 2. We were not able to administer the Cohort 3 teacher surveys in spring 2020 due to COVID-19-related school closures. These Cohort 3 surveys were instead administered (virtually) in fall 2020.

grade year, the school year following the study year.[12] This assessment provided measures of the impact of BARR on 10th-grade students' performance on a standardized test. The specific measures that we derived from these data are described below.

Measures of Ninth-Grade Teachers' Perception of Self, Students, and Schools.
The teacher survey was developed for the i3 validation study and used in that study. It captures nine constructs of teacher experiences and attitudes:

- Strength-based mindset
- Postsecondary educational expectations
- Student academic motivation
- Perception of students' behavior, commitment, and attitudes
- View of students' observed behavior, commitment, and attitudes
- Use of data
- Teacher self-efficacy
- Collaboration with and view of colleagues, and
- View of school supports provided to teachers.

We used Rasch modeling to create scale scores from the survey items for each measure, which were then standardized to have a mean of 50 and a standard deviation of 10. Details about the survey scales, including their reliability, are provided in Appendix A. The survey scales and items are in Appendix B.

Measures of Ninth-Grade Students' Academic Outcomes and Behavior in School. The academic and behavior outcomes measured are as follows:

- **Core course credits earned in ninth grade.**[13] To determine the impact of BARR on credit attainment, we collected transcript data for all Grade 9 students. We transformed these transcript data into a **continuous measure** of the percentage of core credits that students attained during Grade 9. We measured Grade 9 students' average course credit attainment for all core courses (i.e., English language arts [ELA], math, science, social studies) attempted during the school year.

[12] For Cohort 3, school closures severely limited PSAT test administration in fall 2020. Consequently, Cohort 3 schools were excluded from the PSAT analyses.
[13] In Cohort 3, schools had varying policies for whether it was possible for students to lose credit after COVID-19-related school closures. Schools adjusted their policies because some students and teachers had limited access to technology or a stable internet connection at home, and others had limited familiarity with online teaching platforms. As such, only transcript data from the first half of the school year (i.e., through Semester 1, Quarter 2, or Trimester 2) were included in the analysis.

- **Passing all core courses in ninth grade.**[14] To determine the impact of BARR on course failure, we transformed transcript data into a dichotomous measure of course failure (i.e., passing all core courses) during the ninth-grade school year.
- **GPA at the end of ninth grade.**[15] To determine the impact of BARR on GPA, we used transcript data and transformed letter and numeric grades into a standardized 0 to 4.0 grade point scale based on the College Board GPA scale (College Board, 2021). We measured students' average GPA at the end of ninth grade.
- **Chronic absenteeism in ninth grade.**[16] To determine the impact of BARR on chronic absenteeism, we used student attendance data to measure the proportion of Grade 9 students who missed 10% or more of enrolled school days during the school year.
- **Suspensions in ninth grade.** To determine the impact of BARR on suspensions, we used student discipline data to measure the proportion of Grade 9 students who received one or more in-school or out-of-school suspensions during the school year.
- **Persistence to 10th grade.** To determine the impact of BARR on persistence to the next grade, we used students' 10th-grade enrollment data to measure the proportion of Grade 9 students who enrolled in 10th grade in the same school the year following implementation.

The administrative data also provided demographic information about each student, including gender, race/ethnicity, EL status, special education status, and whether students qualified for FRPL. We used these measures as control variables in statistical models and as indicators to identify different groups of students for analysis of how the impact of BARR varied across different populations. Additional details about the data sources and measurement for all outcomes are provided in Appendix A.

Measures of Students' Performance on Standardized Tests. To determine the impact of BARR on academic achievement, study schools administered the PSAT in the fall of students' 10th-grade year.[17] The outcomes measured from this test are as follows:

- **Performance on ELA PSAT/NMSQT in the fall of 10th grade.** For this measure, we used fall Grade 10 students' average PSAT (or similar test) evidence-based reading and writing section score.

[14] The same limitations introduced above about the use of transcript data for Cohort 3 schools apply to this outcome as well.
[15] The same limitations introduced above about the use of transcript data for Cohort 3 schools apply to this outcome as well.
[16] To account for the lack of consistency in Cohort 3 attendance data during the COVID-19 pandemic, we used the proportion of Grade 9 students who missed 10% or more of the school days between the start of the school year and February 28, 2020.
[17] We made some exceptions to this administration window, allowing four schools in Cohort 1 and 2 to administer the PSAT-10 in the spring of 10th grade instead. Because of concerns with the comparability of PSAT-10 scores with the PSAT/NMSQT scores collected in all the other schools, we dropped these four schools from the impact analyses presented in this report. We conducted a sensitivity test with these schools included in the analyses and found that the results were consistent.

- **Performance on Mathematics PSAT/NMSQT in the fall of 10th grade.** For this measure, we used Grade 10 students' average PSAT (or equivalent test) evidence-based mathematics section score.

BARR Scale-Up Measurement

We used program records collected from BARR Center and Hazelden Betty Ford Foundation (a partner of BARR) and interview data collected from BARR regional managers and BARR coaches to understand the facilitators and barriers to successful scale-up of the BARR model during the grant period and the extent to which BARR scale-up strategies were implemented as planned during the grant period (June 2017–May 2021). Measurement was based on predetermined thresholds set for each indicator of BARR scale-up, which were established in consultation with BARR Center. Details about the scale-up data sources and measurement are provided in Appendix G.

Analytic Approach to Impact Analyses

To estimate the impact of BARR, we used statistical models to compare the outcomes of students and teachers in schools that were assigned to the BARR model to students and teachers in the business-as-usual control group. We used ordinary least squares (OLS) regression analysis for continuous outcomes (the percentage of credits attained, GPA, PSAT outcomes, and teacher experiences/attitudes) and logit maximum likelihood analysis for dichotomous outcomes (passing all core courses, chronic absenteeism, suspension, and persistence to 10th grade).

To improve the precision of the impact estimates and address any differences in student and school-level background characteristics between the BARR schools and the control schools, we included student- and school-level control variables (covariates). To account for the blocked random assignment design and for variation in random assignment ratios across blocks, the impact regression models also included dummy variables for each of the study's 30 random assignment blocks.[18] To account for nesting of students within schools, we estimated Huber-White robust standard errors (Zeileis et al., 2020). We converted all estimates into effect sizes using Hedge's g with a small sample size bias correction for continuous outcome measures and Cox's index for dichotomous outcome measures (What Works Clearinghouse, 2020).

[18] These dummy variables mirrored the randomization blocks wherever possible. We used different groupings in two types of cases: (1) there were five schools that were randomized one by one, necessitating them to be grouped with similar schools in other blocks; and (2) there were some small blocks where outcome data were unavailable for all treatment or all control schools in the block. Those schools were then also grouped with another similar block for the purpose of that particular outcome analysis.

In the analysis of administrative and PSAT data, our models included student-level demographic information (race/ethnicity, gender, FRPL status, EL status, and special education status), prior student achievement (student eighth-grade academic performance on state standardized tests in ELA and mathematics), and school-level graduation rates from the year prior to the study year.

The statistical models used in the analysis of teacher survey data included individual-level variables as collected in the survey as well as school-level variables. The control variables used in the model included gender, race/ethnicity, years of experience, and education level. At the school level, we included variables for school size and an indicator for rural schools.

Given the unique conditions affecting each of the study's three cohorts, we estimated all models separately for each cohort. Afterwards, we then used a fixed-effects meta-analytic model to combine effect sizes across cohorts (What Works Clearinghouse, 2020; Hedges & Vevea, 1998). This approach treats each cohort as a distinct study and obtains an overall effect by calculating an average of the three cohorts' effect sizes weighted by the standard error of each cohort's effect size. To obtain estimated regression-adjusted means and impact estimates for the combined cohorts, we calculated weighted averages of the regression-adjusted group means for the three individual cohorts.

School-Level Attrition and WWC Standards (What Works Clearinghouse, 2020)

According to WWC standards, the school-level attrition rates for our three confirmatory academic outcomes (Credits Earned, PSAT ELA, and PSAT Math) align with low levels of potential bias under both optimistic and cautious sets of assumptions at the combined cohort level. For Cohort 1, potential bias from school-level attrition is low only for the Credits Earned outcome. In Cohort 2, the potential bias levels are considered low under optimistic assumptions for each of the three confirmatory outcomes. When looking at student-level attrition within analysis schools, the potential for bias was low for the Credits Earned outcome, but high for PSAT outcomes.

Complier Average Causal Estimate. Three schools assigned to the BARR treatment group in Cohort 1 and one treatment school in Cohort 2 did not implement the BARR model during the study year. These schools are considered noncompliers and are included as BARR treatment schools in all impact analyses. Also, in eight treatment schools, a subset of students was not part of the BARR model. (These schools were "partial implementers" and are also included in all impact analyses.) These schools and students can be considered "crossovers" in an RCT like this evaluation. To account for these crossovers, we estimated supplemental "complier average causal estimates" (CACE; Angrist et al., 1996) to present together with the experimental impact estimates in this report. These CACE estimates are nonexperimental, but they use the underlying experimental design to remove potential selection bias (i.e., underlying differences between complier and noncomplier schools and students). In Chapter 3, where we present impact estimates, we describe the main intent-to-treat impacts, including CACE estimates

where possible, to show what the impacts would have looked like if all schools and students assigned to the BARR group participated in and/or were exposed to the program.

Analyses of Variation in Impacts across Student Groups. We also conducted analyses of variation in impacts across different groups of students (also known as "subgroup analyses") to examine whether program effects varied by student background characteristics, such as gender, race/ethnicity, English learner status, special education status, and free and reduced-price lunch status. Even though the BARR model targets all ninth-grade students, impacts are expected to be stronger for students who traditionally have more difficulty with the transition into ninth grade and have worse academic outcomes, including, for example, lower income students, English learners, and male students. For these analyses, we estimated the impact models separately for different groups of students.

Attrition Analysis. Sample attrition occurs when the final analytic sample differs from the randomly assigned sample due to the loss of participant data. Understanding the extent of attrition is important because the composition of the randomly assigned groups can change when attrition occurs, and this change can lead to bias in the estimated effect of a treatment. We assessed this potential bias due to sample attrition using standards developed by the What Works Clearinghouse (2020).[19] The number and percentage of schools for which we successfully collected outcome data are reported in Exhibit 1.3.

Exhibit 1.3. School Analysis Samples

	Treatment Schools			Control Schools		
	Total *N*	Number	Percent	Total *N*	Number	Percent
Administrative Data						
Cohort 1	15	15	100	7	7	100
Cohort 2	11	10	90.9	11	10	90.9
Cohort 3	11	11	100	11	10	90.9
Total	37	36	97.3	29	27	93.1
PSAT Data						
Cohort 1	15	12	80.0	7	7	100
Cohort 2	11	10	90.9	11	9	81.8
Total	26	22	84.6	18	16	88.9

[19] As part of its evidence screening process, the What Works Clearinghouse (WWC) describes whether the combination of overall attrition (i.e., the percentage of the overall sample that is lost from the initial randomization) and differential attrition (i.e., the difference in the rates of attrition for the treatment and control groups) is high or low, then determines whether the expected bias due to attrition is tolerable or unacceptable (What Works Clearinghouse, 2020). When attrition falls into the low attrition category, the analysis is eligible for the highest rating of *Meets WWC Group Design Standards Without Reservations*. When attrition falls into the high attrition category, the analysis is only eligible to be rated *Meets WWC Group Design Standards With Reservations* if it satisfies the WWC's baseline equivalence requirement.

	Treatment Schools			Control Schools		
	Total *N*	Number	Percent	Total *N*	Number	Percent
Teacher Survey						
Cohort 1	15	14	93.3	7	7	100
Cohort 2	11	10	90.9	11	11	100
Cohort 3	11	10	90.9	11	7	63.6
Total	37	34	91.9	29	25	86.2

Note: Administrative sample numbers are based on the outcome variable capturing the percentage of core credits students earned. Other variables may have slightly more or fewer observations.

The number and percentage of students in schools for which we successfully collected outcome data are reported in Exhibit 1.4.

Exhibit 1.4. Student Sample in Analysis Schools

	Treatment Schools			Control Schools		
	Total *N*	Number	Percent	Total *N*	Number	Percent
Administrative Data						
Cohort 1	4,246	4,060	95.6	2,059	1,943	94.4
Cohort 2	3,534	3,332	94.3	3,864	3,490	90.3
Cohort 3	4,266	4,098	96.1	2,865	2,605	90.9
Total	12,046	11,490	95.4	8,788	8,038	91.5
PSAT Data						
Cohort 1	3,063	2,193	71.6	2,059	1,250	60.7
Cohort 2	2,752	1,777	64.6	3,756	2,052	54.6
Total	5,815	3,970	68.3	5,815	3,302	56.8

NOTE: The "Total N" column is for schools included in the analysis. Administrative sample numbers are for the Credits Earned variable. Other administrative outcome variables may have slightly more or fewer observations.

The number and percentage of teachers in study schools who responded to the teacher survey are reported by cohort in Exhibit 1.5.

Exhibit 1.5. Teacher Sample in Analysis Schools

	Treatment Schools			Control Schools		
	Total *N*	Number	Percent	Total *N*	Number	Percent
Teacher Survey						
Cohort 1	114	81	71.1	90	65	72.2
Cohort 2	162	102	63.0	158	100	63.3
Cohort 3	146	93	63.7	101	65	65.4
Total	422	276	65.4	349	230	65.9

Note: The numbers are based on the Academic Motivation variable. Other variables may have slightly more or fewer observations.

Baseline Equivalence. To assess the baseline equivalence of the treatment and control groups, we compared baseline and demographic variables for schools in the analytic sample. We found no statistically significant differences between the treatment and control schools on prior graduation rate which is aligned to our academic outcome measures. While there were some differences within a given cohort between treatment and control schools greater than a 0.25 standardized difference (the WWC range for acceptable differences requiring statistical adjustment), none were statistically significant (Exhibit 1.6). Additionally, demographics across the three cohorts were quite similar. Appendix Exhibits A6.1-A7.1 show the baseline sample means for the PSAT and teacher survey data.

Exhibit 1.6. Baseline Equivalence for Study Schools in the Administrative Data

	Treatment Mean	Control Mean	Standardized Difference	*P*-Value
Cohort 1				
Graduation rate	85.7	90.2	-0.48	0.243
School demographics				
Percent female	48.5	47.6	0.26	0.592
Percent students of color	57.9	58.5	-0.02	0.974
Percent FRPL	62.4	50.2	0.58	0.318
Percent EL	7.2	8.3	-0.16	0.746
Percent special education	15.7	17.0	-0.25	0.492
Total schools	15	7		
Total students	4,060	1,943		
Cohort 2				
Graduation rate	82.0	80.6	0.16	0.731
School demographics				
Percent female	47.6	48.1	-0.14	0.753
Percent students of color	81.7	75.3	0.23	0.621
Percent FRPL	68.4	74.5	-0.28	0.555
Percent EL	23.3	28.1	-0.23	0.620
Percent special education	14.7	15.9	-0.21	0.654
Total schools	10	10		
Total students	3,332	3,490		
Cohort 3				
Graduation rate	88.2	87.4	0.09	0.845
School demographics				
Percent female	48.8	47.7	0.37	0.412
Percent students of color	61.6	61.4	0.01	0.989

(continues)

	Treatment Mean	Control Mean	Standardized Difference	*P*-Value
School demographics (cont.)				
Percent EL	13.7	7.9	0.54	0.235
Percent special education	15.4	14.3	0.17	0.710
Total schools	11	10		
Total students	4,098	2,605		
Cohorts 1, 2, and 3 Combined				
Graduation rate	85.5	85.1	0.04	0.873
School demographics				
Percent female	48.3	47.8	0.15	0.563
Percent students of color	66.2	66.8	-0.02	0.933
Percent FRPL	66.6	60.4	0.26	0.360
Percent EL	14.4	16.9	-0.15	0.578
Percent special education	15.3	15.6	-0.07	0.789
Total schools	36	27		
Total students	11,490	8,038		

Note: FRPL = free or reduced-price lunch; EL = English learner.
Source: Graduation rates obtained from the US Department of Education's EDFacts. American Institutes for Research administrative data. The numbers are based on the Credits Earned variable and include students in the earlier joiner sample frame.

Additional details about the analytic approach are provided in Appendix A.

Chapter 2. Implementing BARR in the Study Schools

There are two primary data sources supporting the implementation research for this evaluation, as described in Chapter 1. The first data source is a series of interviews we conducted with teachers, administrators, and staff from 32 BARR schools[20] after their first year of implementing the BARR model. The second data source is a series of ratings by AIR researchers and school BARR coordinators that together enable us to create a fidelity of implementation measure for each of the eight BARR "structural" components and for the model as a whole. Exhibit 2.1 provides a detailed description of each of the eight components that we explore in this chapter.

Exhibit 2.1. Description of the Eight Components of the BARR Model

Component 1: Focus on the Whole Student.
The BARR model explicitly focuses on the whole student, not just on a student's performance in a particular subject or his or her specific academic or nonacademic challenges. Thus, teachers and administrators are instructed to identify each student's assets and leverage them in addressing challenges and barriers. Working across multiple core courses is intended to make it easier to identify these assets and to address challenges that manifest themselves differently in different settings. Teachers also learn about their students' extracurricular interests and strengths (e.g., sports, friendships, passions) so they can reference and leverage them to support the student's academic performance and engagement in their classroom.
Component 2: Provide Professional Development for Teachers, Counselors, and Administrators.
Teachers and school administrators receive hands-on training and coaching to improve their communication with other school staff about students' progress, assets, and barriers as well as their ability to identify and implement necessary interventions to help keep students on track. This includes BARR staff and coaches directly observing block/team meetings, including teachers and BARR coordinators, and providing feedback on the tone and pacing of these meetings, the solutions and interventions being developed, and the degree to which these interactions are consistent with best practices developed and identified in other BARR settings.

(continues)

[20] As noted earlier, the study included three cohorts of schools. Cohort 1 had 15 schools assigned to treatment. However, three Cohort 1 schools did not implement BARR and are not included in our interview data. Cohort 2 had 11 schools assigned to treatment. However, one Cohort 2 school did not implement BARR and is not included in our interview data. Cohort 3 had 11 schools assigned to treatment. However, there was one school where we were not able to conduct interviews with staff due to delays in approval related to COVID-19.

Exhibit 2.1. Description of the Eight Components of the BARR Model (continued)

Component 3: Use BARR's I-Time Curriculum to Foster a Climate of Learning.
The BARR model includes a weekly I-Time lesson, which is taught (in a weekly rotation) by one of the core subject teachers and explicitly addresses students' social and emotional development and related issues. The I-Time activities specifically aim to improve student-to-student and student-to-teacher communication and to support mutual understanding and collaboration. The sharing of personal experiences and beliefs is a major part of these activities, which enable students and teachers to better understand and appreciate one another's circumstances and motivations.
Component 4: Create Groups of Students with Common Teachers.
The course schedule is restructured such that distinct groups of students share the same group of teachers for their core subjects. This structure is intended to increase feelings of community and belonging among students and enables their teachers to compare and improve students' academic progress across the different subjects.
Component 5: Hold Regular Block/Team Meetings.
A key feature of the BARR model is weekly block/team meetings during which the core subject teachers and the BARR coordinator discuss the academic progress, assets, and needs of each student. During these meetings, the team agrees upon any interventions that individual students may need and who will take responsibility for implementing them. The implementation and effectiveness of these interventions are discussed and monitored in subsequent meetings.
Component 6: Conduct Risk-Review Meetings.
Students who persistently fail or exhibit major attendance or behavioral challenges or who experience serious problems outside of school (e.g., food insecurity, addiction issues, homelessness) are referred to risk review meetings, which include school counselors, school administrators, and other support staff, such as community mental health counselors or school resource officers. These meetings identify specific interventions, and the implementation and success of these interventions are monitored in subsequent meetings. Those interventions include referrals to connect students and their families to community resources.
Component 7: Engage Families in Student Learning.
The BARR model seeks to support ongoing interaction with families to ensure their continued engagement in their child's education. Many of the interventions agreed upon in block/team meetings and risk review meetings include family outreach components. A similar, assets-first approach that is used in these meetings is applied to conversations with families.
Component 8: Engage Administrators.
The BARR model requires ongoing commitment from school leadership (e.g., time, attention, staff resources). BARR staff aim to directly involve school administrators in the day-to-day implementation of the model. In addition to ensuring organizational support, such engagement seeks to enhance administrators' ability to make decisions, support their teaching staff, and take an active role in their students' academic and nonacademic success.

We begin this chapter with a description of the Year 1 implementation from the perspective of school staff. The findings are organized by the eight BARR components and highlight facilitators of and barriers to successful implementation of the BARR model in their schools. In the second half of the chapter, we present measures of fidelity of implementation, which will inform the interpretation of the results of the impact analyses we present in subsequent chapters.

BARR Implementation from the Perspective of School Staff and Leaders

Professional Development

BARR schools receive professional development and coaching to help them implement and sustain the BARR model throughout their first year of implementation and beyond. These foundational training and BARR coaching visits were the professional development activities respondents most frequently cited as facilitators of implementing the BARR model. All BARR schools across cohorts reported that all or most of their BARR teachers and staff involved in BARR (e.g., counselors, administrators) attended the foundational training held on BARR model theory and practice. Some respondents who participated in the training described it as "*excellent*" and "*exceptional.*" An educator from one school shared:

> The BARR training that we had before the beginning of the school year was excellent. And that was my first experience as they walked us through the activities by actually participating in them, and that was the first time I thought, 'This is different. This is much better, and this is going to make a big difference.'

Similarly, respondents from nearly all BARR schools shared that BARR coaching visits provided schools with actionable feedback informed by observations of meetings and I-Time lessons. A school leader described their experience:

> The coaching visits are what make real time differences in our program because we have one-to-one help at our school about our teachers and our program and our students. [With] our coach here on site, we were able to go deeply into our strengths and our needs. And she participated [in] and observed risk review meetings, I-Time and was able to give very specific, targeted feedback and support to our team. So, by far the coaching visits had the biggest impact in our implementation.

When asked about barriers to successful professional development, some BARR schools reported that scheduling professional development sessions was a challenge. For example, about a third of schools in cohorts 1 and 2 reported experiencing difficulties scheduling monthly meetings consistently due to competing events, such as holidays, testing windows, and instructional meetings. A BARR coordinator in one school shared:

> The way our school runs sometimes, that meeting was taken away from us and I wasn't able to have it on that particular Monday even though I had scheduled to have it biweekly. They were my honored times, but sometimes that time wasn't honored due to other factors in the school.

Despite the scheduling challenges that some schools experienced, our fidelity of implementation analysis (detailed in the second half of this chapter) mirrored the largely positive experiences with BARR professional development and coaching that staff reported in their interviews with us. That is, professional development was one of the components that schools were most often able to successfully implement with fidelity.

Restructuring

The BARR model requires schools to restructure their class schedule in ways that impact students, teachers, support staff, and administrators. This is to ensure that groups of students share the same core subject teachers, so that they can then work together to better support their students across all core subjects. In addition, schools are expected to appoint a BARR coordinator, who is the school's point person in its interactions with the BARR Center and who also is responsible for coordinating BARR activities in the school.

Interview data collected across each study cohort described whether and how BARR schools restructured their ninth-grade schedule to implement the BARR model. Nearly all schools across the three cohorts reported restructuring their ninth grade so that at least 80% of students had the same shared teachers in at least three core classes.[21]

In general, BARR schools spoke favorably about the restructuring. For example, a teacher said:

> We all as teachers, have those same students and so we can compare our experiences… We get so much more valuable data and observations, and we can compare notes, and I just, I don't want to teach without it. It's the most supportive for students and for teacher morale and cooperation.

Another teacher reported that the restructuring had *"helped it be very team orientated,"* adding that the BARR team *"by far, has been one of the best grade level teams that I've ever been on."*

However, schools were not given guidance on how to group students, and there was considerable variability in what factors schools considered. Some schools grouped students at

[21] Although 11 of the 12 Cohort 1 schools indicated restructuring their ninth grade so that at least 80% of their students have the same shared teachers in at least three core classes, five of the 12 Cohort 1 schools implemented BARR with only a subset of their ninth-grade students.

random while others based it on an indicator of academic performance, such as achievement in math and student selection of honors classes or academies. About a third of BARR schools reported experiencing challenges related to how the students were grouped, mostly related to how honors courses were handled. For example, a BARR teacher commented that excluding honors students from BARR was *"a little disastrous"* because it resulted in the BARR cohort of students having:

> Some fantastic, high achieving students mixed in with an overwhelming number of low achieving students. And if you had a smattering of low achieving students you could pair them up with somebody that's more skilled, and probably go at a little higher pace than we've been able to do this year.

A teacher at another BARR school shared, *"If a kid selected one honors course, then they put them in all honors courses so that they would share core teachers. But then those students struggled by having a full course load of honors classes."*

A few schools also noted that they had trouble developing a schedule that would allow 80% of their students to have the same teachers within a cohort due to student population factors (e.g., having a transient student population) or coordinating schedules for teachers who teach grades other than ninth grade.

As part of implementing BARR in a school, BARR encourages that a school counselor attend BARR meetings and activities. Across cohorts, about three quarters of BARR coordinators reported in their self-assessments that counselor involvement with BARR was in place. The remaining schools reported that a counselor was involved but not in a consistent manner, or that they experienced challenges with their counselor's involvement with the model. One school explained, *"We did have counselors assigned to [BARR] teams, but they also are assigned to all other grade levels. So, it was a little difficult sometimes to get counselors there for every single [BARR] meeting, but they were involved."*

Schools varied in whether they conducted block meetings and risk review meetings weekly. About three fourths of schools held block meetings weekly, and slightly fewer, around two thirds, conducted risk review meetings weekly. A small number of the remaining schools reported conducting meetings at least once a month; only one school reported that they did not conduct block meetings at all. Additional findings on the experience of these meetings are discussed in the block and risk review section below.

Another aspect of restructuring is asking students to complete a questionnaire on their interests and potential mental health, social service, or substance abuse needs. More than three fourths of BARR schools reported administering this questionnaire in their first year of implementation, but

schools varied in whether and how they used the resulting survey data. A total of only five schools across all cohorts explicitly shared *how* they used the student questionnaire data to inform their work and supports for students. One school leveraged the questionnaire data during block meetings "*to get some background on the student, and kind of see what's going on and what supports we could provide for them,"* which they found to be *"very valuable."*

Family Involvement

Almost half of the BARR schools reported conducting a parent orientation that, at a minimum, introduced BARR content. In addition, BARR coaches encouraged schools to establish a Parent Advisory Group, which many schools did, often with support from the coaches.

Teachers and staff at BARR schools communicated with families most frequently via e-mail and phone calls. All BARR schools consistently used phone calls to engage families, and nearly all BARR schools indicated they e-mailed families frequently. In fact, nearly half of the BARR schools shared that the structure of the BARR model made it easier for them to communicate with families more frequently. This was largely due to the collaboration between BARR teachers and staff. A BARR teacher explained:

> We tend to communicate more with parents this year, but it also makes it easier because we can be in a BARR meeting, talking about a student and then decide right then and there well as a team, let's contact this parent. I think that helps encourage more parental involvement.

The BARR structure also facilitated teachers and staff meeting with family members when needed. "*We've definitely been able to get more parents to come to the school,"* a respondent said, adding that they use BARR meeting time to meet with families: "*Having the time to meet on Tuesday …that's when I get teachers and parents up here."* The BARR model's emphasis on highlighting students' strengths and positive qualities also supported BARR schools' family involvement. Teachers from over half of the BARR schools discussed how this practice helped promote parent-teacher relationships. A teacher shared how "*starting off with a positive"* helped them "*gain a lot of parent's trust and understanding," adding:*

> The big piece of BARR that was very beneficial to gaining that parent trust was the first thing we went over always with every parent was the student strengths. Right there they didn't get a bad vibe like, 'Oh you're just out to get my student.' It was like, 'Oh, you see some positive things in my student,' and those things we want to change. I think establishing that trust, not only with the student but with the parent is the same thing. It's very powerful.

Although family involvement was generally discussed positively in BARR schools, some schools reported barriers to successful engagement. Nearly a quarter of BARR schools reported having insufficient time to reach out to family members. One BARR teacher explained:

> Other than seeing them at open house and teacher conferences, it's a lot harder to connect with parents at this level. In the past, I have been able to…invite [them to a] poetry reading, and bring family in, but this year I wasn't able to do that, because there were limitations of time… So sadly, I think I've declined on my parent connections.

Despite this, the teacher added that they still *"do make sure to contact [families] as much as possible, whether positive or if a student needs a little extra help."*

Although close to half of the BARR schools reported that the structure of the BARR model made it easier for them to communicate with families more frequently, we found that the generally positive experiences that staff described overall in interviews did not translate into high levels of fidelity of implementation for the family involvement component, as detailed in the second half of this chapter. Family involvement was one of the components that schools most struggled to implement in the first year of BARR implementation.

I-Time Curriculum

BARR's I-Time is a weekly teacher-led activity that creates social-emotional learning experiences for students. I-Time "softens" the teacher-student interface, making it easier for each person to approach the other about issues that are barriers to school motivation. I-Time also builds teacher and peer awareness of student strengths that can become the basis for positive communication about academic achievement.

Teachers from nearly all BARR schools reported conducting I-Time lessons where at least some students shared information about themselves, took risks, and engaged with teachers. A BARR teacher recounted their last I-Time lesson:

> There were four or five [students] that I was pretty sure weren't going to say a word, but they all shared. And even in…my really hard group, they all shared. It was squirrely a little bit, some of it was a little bit silly, but some kids were downright emotional, and no one burst into tears or crawled into a little ball, but it was a willingness to open up.

One way that BARR teachers created a culture of openness was by sharing their own stories and personal experiences. Respondents from about three fourths of BARR schools reported that it was important to share personal stories during I-Time lessons to help teachers build connections with their students. A BARR teacher explained:

> *One of the most important pieces of the I-Time is basically giving the example of the personal side of it. When [students] see that and they see 'oh wow, not only are you a teacher but you're just a normal person like the rest of us,' that wall breaks down. Kids then start to open up.*

BARR schools reported variation in how I-Time lessons were implemented across and within BARR schools. Although many BARR schools described high levels of student engagement with I-Time lessons, respondents from about a third of BARR schools reported that some students did not participate in discussions, did not share, or felt the activities were not engaging. There also was variation in how schools scheduled I-Time lessons. Respondents from about a third of BARR schools reported feeling they were unable to stay on schedule due to issues such as assessments and teachers forgetting. Teachers are encouraged to adapt I-Time lessons to meet the needs of the students, and respondents from just over half of the BARR schools reported that teachers adapted I-Time lessons, and when they did, I-Time lessons were successful. A BARR coordinator at one school explained:

> *The [I-Time] lessons are valuable. I think, especially if teachers are able to adapt them in just enough of a way to really suit their particular clientele. I feel like in some cases, you can look at a little bit of it and say, you know, I'm not so sure that's going to work for this group. I'm going to change it up a little bit. So, with a strong staff who have that flexibility and understanding of their students, they can be very successful.*

However, about a quarter of BARR schools reported that many of their teachers do not adapt the I-Time lesson. A BARR teacher provided one possible explanation: "*What if a first year teacher has to teach [an I-Time] lesson? It could be totally ineffective if you don't know how to adapt the lesson."* In addition, some I-Time topics were challenging for teachers who were not hired or trained to deliver a social-emotional curriculum, as a BARR teacher explained:

> *When we got a lesson on grief, I had a counselor come in here... There are certain things that are out of my wheelhouse. I'm just not going to talk about how you cope with the death of your mother. I understand, but at the same time, I don't want to be that. That's not what I went to school for. It comes with the territory but there are people on staff who know it better than me so let them come in and talk.*

This speaks to the importance of having a counselor involved in BARR implementation beyond the block and risk review meetings. BARR actively encourages such involvement, especially in sensitive I-Time lessons like the one focused on grief referenced above.

Block and Risk Review Meetings

Key features of the BARR model are regular block meetings and risk review meetings. During weekly block meetings, the core subject teachers and the BARR coordinator discuss the academic progress, assets, and challenges of each student. Students identified as high risk in

block/team meetings are elevated to risk review meetings, which also occur weekly. Risk review members include school counselors, school administrators, and other support staff, such as community mental health counselors or school resource officers who identify specific interventions and monitor their implementation. BARR teachers are expected to review progress of all students with whom they are blocked, and teachers, administrators, and support staff who participate in the risk review are expected to review progress of students identified as at risk by the block team(s).

Our data collection interview protocol did not distinguish between block and risk review meetings. Instead, it asked about "team meetings" generally with the goal of allowing respondents to share about both meeting types. Respondents often described block and risk review meetings in a way that we could not clearly differentiate from each other. As such, implementation of block and risk review meetings are discussed together in this section.

When we asked BARR teachers whether there were any particular aspects of the team meetings that they thought helped them as a teacher, some of the most common responses were about the meetings' structure, teacher collaboration around students, and the staff who participated in the meetings. Staff in about three quarters of BARR schools reported that following a structured process for conducting block and risk review meetings facilitated the meetings' success. A respondent praised the materials used during the meetings: "*I think that's one thing I love about BARR. I love every protocol for each section."*

Another described how the block meetings were engaging and inclusive:

> "I feel that every teacher has that moment to voice their opinion and that's a powerful piece."

Three quarters of BARR schools indicated that BARR teachers and staff shared their knowledge of and experiences with students to contribute to a greater understanding of the individual student's status and progress. Although this was not exclusive to BARR meetings, it was most frequently discussed as occurring during block and/or risk review meetings. One BARR teacher described the experience of collaborating with other BARR teachers who had the same students as "*the biggest positive impact the BARR [model] has, from a teaching standpoint."* In addition, another BARR teacher described why it was beneficial to include administrators and support staff in risk review meetings:

> Our administrator and our counselor join the [risk review] meeting. Now they can add additional things… Have they had contact with the parents? Have the parents contacted the school? Has the student been in the office? Has the student been sent to the counselor for anything? Because sometimes they have information that we don't have access to.

BARR schools also reported challenges and barriers to implementing the block and/or risk review meetings with fidelity. About three quarters of BARR schools indicated that insufficient time was a challenge to implementing block and risk review meetings. Sometimes there was not enough time to conduct the meeting or to get through the whole student caseload for a meeting. A respondent described how the meeting time was often insufficient: "*We tend to go over time, or we don't get to everything because we're so focused on one student. So that is challenging because it feels like we never complete an agenda of things."*

Another respondent felt the meeting time was sufficient, but the frequency of meetings was cumbersome:

> I would say the most challenging thing, [is] not with the meetings themselves, but the frequency of meetings at times. I understand that nobody calls a meeting for no reason. But there are times where students may be having a crisis and we meet four times a week instead of two.

Three additional aspects of the block and risk review were experienced as challenges by about half of the BARR schools: identifying and discussing students' strengths, identifying goals for students, and selecting and monitoring interventions. About half of the BARR schools indicated that it was a challenge to identify student strengths. This includes a tendency to skip the identification of student strengths or focus disproportionately on students' negative behavior:

> You try to be positive about every kid, but the children that you're most focused on are the children who are really struggling or really problematic. So, you know, when you put up a name, everybody has something to say about that kid. So, you don't always want it to turn into a giant list of complaints because, you know, everybody has some redeeming quality… And then we often times neglect to mention that positive aspect. What that kid has that maybe no other kid does. So, that's the hardest problem I think with those meetings is not turning it into a giant list of complaints.

About half of BARR schools indicated that they experienced challenges with setting specific goals for students. "*I think our goals were very general,"* a BARR coordinator shared. "*Our coach was helping us out to be a little bit more specific with our goals."* Finally, about half of BARR schools indicated that they experienced challenges with intervention selection for the hardest-to-help students:

> The most challenging part, I think, would be not feeling like there's solutions…There've been a handful of students where we've tried lists of different accommodations and supports for, and it seems like nothing's working [for] them. And we're still concerned for them, so we still put them on our list, but we don't know what to do.

In the interviews, BARR teachers reported experiencing similar levels of successes and challenges for the block and risk review meetings. However, when we examined the fidelity of implementation data, which we detail below, we found that the implementation of these two types of meetings varied.

Whole Student Emphasis

The BARR model explicitly focuses on the whole student, not just on a student's performance in a particular subject or their specific academic or nonacademic challenges. The whole student emphasis component of the BARR model addresses the extent to which teachers worked collaboratively to develop an understanding of educating the whole student.

BARR coordinators described the extent to which their teachers understood the whole student concept and expanded their perception of their role as a teacher. About half of BARR coordinators reported that their teachers understood the whole student concept. One BARR coordinator commented:

> I think we have a lot of students who have connected deeply to our teachers. The students share openly with the teachers, and our teachers do a great job of being there before school, after school. Really supporting them not just in school stuff but in outside of school stuff as well.

About two thirds of BARR schools reported that teachers successfully broadened their perception of their role from purely providing academic instruction to supporting all facets of students, from academic to nonacademic strengths and needs. At least a few schools in all three waves attributed this change directly to the BARR model:

> One of the things that BARR really emphasizes that has helped many of our teachers is relationships. I think because BARR's core really states that relationships first and then other things will follow. I think that that emphasis has really helped teachers to broaden their perception of their role as a teacher. That's not just about instructing content, that it really does start with the relationship.

In about half of BARR schools, some teachers reported confidence and competence in working with the whole student while other BARR teachers did not. Speaking about the teachers' confidence level, one BARR coordinator said, *"Some teachers are really good at it. Other teachers are still learning and perfecting working with the whole student."* Another BARR coordinator spoke about a challenge that teachers faced when trying to work with the whole student: "*There are some students that are dealing with things that the teachers just aren't well versed in…that's not their area. In those cases, it's hard to figure out what to do."* Finally, yet another BARR coordinator shared how teachers are growing in this area:

> I think that the teachers are doing better with that, especially through BARR and starting to really look at all the different factors instead of just academics. Everything that the [BARR] program brings is, I think, helping our teachers look further past the academics.

Contextual Support

The BARR model requires ongoing commitment from school and district leadership (e.g., time, attention, staff resources). Contextual support refers to the extent to which the superintendent, principal, BARR coordinator, and school board are knowledgeable and supportive of the BARR model.

In general, school staff and leaders perceived there to be adequate levels of support for the BARR model at their schools. Respondents from nearly all BARR schools reported that school administrators attended BARR block and/or risk review meetings (at least as often as or for as long as they could) and/or met with the BARR coordinator regularly to discuss the model's implementation. A respondent shared:

> The associate principal … makes sure that people feel supported, and her goal is that ultimately kids are supported, so she's present in a lot of our team meetings, follows up with us to see if kids are improving, or if they need more help.

Still, some respondents from about a third of BARR schools commented that their principal attended meetings less often than what is recommended by the BARR model.[22] This was frustrating for BARR teachers, but did not necessarily mean that school leaders were not supportive of the BARR model:

> Our BARR coordinator told us that the principal needed to be involved in those meetings and he was often absent quite frequently which was a little frustrating for us. Not that he wasn't supportive. He really allowed us the freedom with the BARR program to do as we saw fit.

Respondents also reported being supported outside of team meetings. About half of BARR schools shared that their school leadership demonstrated a willingness to help teachers in general, specifically with BARR implementation, as a BARR teacher described:

> I mean the administration here is really willing to help out and come in your classroom if you need it, or whatever you need. Especially as a BARR team, it being the first year of having BARR here, we're able to take some liberties in terms of doing a BARR retreat for

[22] We interviewed more than one educator at each school, and their responses often varied due to different experiences within the school. For example, within the same school, there could be staff who reported administrators attending meetings regularly, and other staff who indicated that administrators did not attend meetings regularly.

> just the teachers. They're supportive with getting us subs, and just allowing that to happen. So, yeah, the administration has been really helpful just in terms of whatever we need, they just do their best to make it happen.

BARR coordinators also played a pivotal role in supporting BARR implementation. Respondents from about two thirds of BARR schools reported that their BARR coordinator supported implementation in multiple, significant ways, including organizing meetings, collecting data and supplies that teachers needed, and drawing support for BARR from the administration. A BARR teacher explained:

> [The BARR coordinator is] the one who assigned us all what to do for our meetings and I-Time to make sure we don't go over our allotted time for every student because we try to…spend eight or nine minutes per student so we get through the list. She did coordinate in that way. She's also coordinated when our meetings are, what we're going to do. She coordinates who's teaching what I-Time, what our rotations are.

Another BARR teacher shared how the BARR coordinator supported the model:

> [The BARR coordinator has] been the driving force behind [implementing BARR] at the top level. She's drawn support from the other admins, but she's been the person on the admin side that's done a tremendous amount of work for us in so many ways.

Implementation During COVID-19

In addition to conducting a holistic analysis of BARR school interview data to identify the facilitators and barriers of implementation across cohorts, we also analyzed Cohort 3 interview data to understand how COVID-19 affected the Cohort 3 BARR schools' first year of implementation. The campuses of all Cohort 3 schools closed in spring 2020 due to COVID-19. Due to these unique circumstances, we conducted phone interviews instead of site visits with BARR coordinators, teachers, and administrators in fall 2020. We modified the interview protocols from previous years to capture the unique circumstances brought upon by COVID-19 and asked respondents to reflect on the 2019–20 school year, during which school campus closures initially took place.

We found that schools approached BARR implementation in different ways during school campus closures. Overall, respondents from Cohort 3 BARR schools indicated that school campus closures affected whether and how they implemented the structural components of BARR and shared how their experience with this model supported their work with students and families during this period of remote learning.

About half of Cohort 3 BARR schools were able to provide I-Time lessons virtually during the spring 2020 campus closures, with support from the BARR Center, which converted all I-Times into a virtual format. Among the schools that continued to provide I-Time lessons, some noted that their BARR coordinator facilitated lessons or that the lessons were optional for teachers to conduct. Some schools also described how they adapted I-Time lessons for remote learning. Respondents from schools that were unable to continue offering I-Time lessons during campus closures noted overall changes in school requirements, which affected participation in the lessons. One teacher elaborated:

> Our district went full throttle on the closure. They told the kids nothing from the rest of the year was going to count...So again, my classes, it's kind of hard to do an I-Time with two people. You need to have a group.

Furthermore, there was variability in whether and how schools conducted BARR meetings—block meetings and risk review meetings—during school campus closures. The BARR Center provided support for converting these meetings into virtual meetings. The BARR team at one school was unable to meet just prior to and during campus closures due to personnel issues outside of BARR and COVID-19, but, as a school, staff across different grade levels continued to refer students to supports and services in the absence of the meetings. At all the other schools, BARR meetings continued but varied in frequency, duration, and/or format, depending on their specific circumstances and schedules. One teacher elaborated on how their school adapted their style of meetings to support student needs during campus closure:

> …Because [of] the cohort group that was working together, the teachers that were working and the team members, there was this informal sharing [of] information…[on] students that were going missing after the shutdown…And so, it wasn't formal as we had done in the past, but it was a lot of the same sharing of information and making sure that everyone was on the same page when it came to student needs.

Spring 2020 Professional Learning Community (PLC) Case Study

The purpose of the spring 2020 PLC case study was to capture how BARR and the schools they work with were able to continue to collaborate and address the BARR model goals of building intentional staff-to-staff, staff-to-student, and student-to-student relationships while school campuses were closed in spring 2020. The PLC case study provides context for understanding the academic achievement and student engagement results of the impact analysis of the evaluation. Details can be found in Appendix C.

Due to COVID-19, the annual BARR conference was canceled. Schools had the option of having a virtual fourth coaching visit or waiving this coaching session entirely, depending on what would best meet their school's needs. About two thirds of schools indicated that they did not

participate in any BARR training during the spring 2020 campus closures, although the BARR Center offered virtual training and coaching and many schools participated in it. Although schools varied in whether and how they interacted with BARR Center during campus closures, nearly all Cohort 3 schools indicated that the BARR model helped teachers and staff support students during school campus closure. One teacher elaborated:

> Not being able to talk in-person, face-to-face [during campus closure] did have an impact on [student-teacher communication], but at the same time, if it wasn't for us having that prior bond or that prior connection, I think that a lot of students would have just suffered in silence.

Half of the schools indicated that their level of family involvement was low or lower than it had been prior to campus closures; about a quarter of schools reported that the relationships they developed with families through BARR helped maintain communication during campus closures. One teacher shared:

> I felt that since I had spoken to most of my BARR parents anyways, that they were super quick to reach out to me to ask general questions about what was going on with school. "How can my kid navigate this?" So, I felt like they had a comfort level or a touch point because I had already reached out to them as a teacher.

BARR Professional Learning Community Supports During COVID-19

During the school campus closures, BARR Center staff changed their professional development support from in-school to virtual. As part of the i3 scale-up grant, BARR Center staff had already been working to develop and implement video coaching protocols, create new BARR training videos, and build a web-based platform to connect their growing number of schools and individual staff virtually. After the school campus closures, BARR Center staff sought to maintain their focus on the eight components at the core of their model and to maintain fidelity where possible while leaning on the flexible side of their firm but flexible approach. This resulted in a decision to increase the frequency of monthly virtual professional learning community (PLC) sessions to weekly, with each session focusing on a different component of BARR with an emphasis on building connections and community throughout. The BARR coordinator PLCs were BARR's primary mechanism for reaching schools as school campuses were closing.

At the same time, we were exploring alternative ways to continue to evaluate the BARR model's efforts to support schools and schools' efforts to implement BARR considering the campus closures. When BARR Center staff shared that their work with schools was more in demand than before school campus closures and that they were seeing high attendance at the

weekly PLCs, we decided to observe the PLC sessions to capture the work of BARR and BARR schools during school campus closures. We found that the PLCs created an opportunity for BARR staff to (1) create space for adult connection during a turbulent time, and (2) build educator capacity to support students' social and emotional learning and stay connected with students and families while school campuses were closed.

Creating Space for Adult Connection. Creating a space for educators across the BARR network to come together to connect and listen to each other was one of the primary foci of the PLCs that we observed. BARR staff provided opportunities for participants to introduce themselves and respond to an icebreaker question at the start of each PLC, thereby facilitating connections with and among BARR PLC participants from across the country. BARR staff also facilitated connections among participants by having school-based BARR coordinators present at PLCs alongside BARR Center coaches. These presentations created an opportunity for the other BARR coordinators to hear how their peers in other schools were experiencing and addressing the unprecedented issues that schools faced in spring 2020. To engage as many participants as possible, BARR staff relied on breakout sessions to build a forum for connection and listening. BARR staff posed discussion questions about the PLC topic to promote dialogue among participants in the breakout rooms.

Building Educator Capacity to Support Students' Social and Emotional Learning and to Stay Connected with Students and Families. Building educator capacity to support students' social and emotional learning and to connect with students and families was another area of emphasis that we observed in the spring 2020 PLCs. BARR Center staff created PLC meetings around related topics, such as how to adapt the I-Time social-emotional learning curriculum to the virtual environment and how to connect with students, families, and communities after the sudden closure of school campuses. BARR staff also encouraged participants to share successful strategies for building connections and relationships with students and their families. Another way that BARR staff worked to build the capacity of the PLC participants to connect with students, families, and partners was through developing and showcasing tools in collaboration with BARR coordinators, such as a tool to track outreach and contact with students, and a transitions tool. BARR Center also provided specialized professional development on topics such as "How to engage students virtually," through the PLCs by bringing in experts such as a speaker coach.

As part of the pivot, we met with BARR Center staff to discuss the impetus behind their decision to create weekly PLCs as well as the rationale for PLC content and structure. BARR Center staff explained that they intentionally focused on the importance of relationships and trust, and the tone of communication to ensure that, through the PLCs, the schools felt supported and knew that BARR Center was flexible and available to maintain relationships in a virtual format when

schools were ready. From BARR Center's perspective, this was particularly important during this period because educators felt as though they had been given a new and difficult task—educating their students outside the classroom setting—without much direction or knowledge about how to proceed. Appendix C provides more details about the PLC case study.

Measuring Fidelity of Implementation for Study Schools Versus Measuring Sustainable School-Level Implementation for Scale-Up

As part of the fidelity of implementation analysis, we studied treatment schools' first year of BARR implementation. We calculated fidelity using implementation scores collected by AIR researchers during site visits to treatment schools in the spring of their first year of implementation. This differs from the fidelity of implementation analysis conducted for scale-up measurement in Chapter 4, which focuses on BARR Center's longitudinal progress on school-level implementation for all schools that participated in the study.

To What Extent Were the Eight BARR Components Implemented as Intended?

As stated previously, we analyzed the fidelity of BARR model implementation at treatment schools and examined the extent to which schools were able to adequately implement the eight BARR components during their first year of implementation. At each treatment school, AIR researchers observed the implementation of three components (i.e., block meetings, I-Time lessons, and risk review meetings) using fidelity rubrics and interviewed the BARR coordinator to get their ratings on several indicators related to the implementation of the eight key components of the BARR model. We averaged indicator ratings from BARR coordinators and AIR observers to create component scores for I-Time, block/team meetings, and risk review meetings.[23] For the remaining components, only BARR coordinator indicator ratings were used to create each component score. Component scores provided a quantitative measure of the extent to which schools implemented the BARR model. For each school, component scores were compared against predetermined component score thresholds to determine whether schools met model implementation expectations. Thresholds were determined in consultation with the model developers and were based on theory and model implementation practices during the grant period. If a school's component score matched or exceeded the threshold, implementation of the component was deemed adequate. If the component score did not meet or exceed the threshold, implementation was deemed inadequate.

[23] Because of school closures due to COVID-19, we were unable to conduct site visit observations of BARR activities in Cohort 3. We measured implementation fidelity for Cohort 3 using ratings from interviews conducted with BARR coordinators in fall 2020 and the most recent implementation fidelity scores collected by BARR staff during the 2019–20 coaching site visits. For Cohort 1 and 2, we examined the fidelity scores collected by BARR staff and compared them to the scores collected by AIR researchers. We found that the fidelity scores collected by AIR tended to be slightly higher than the scores collected by BARR staff. Because the Cohort 3 fidelity analysis was calculated with scores collected by BARR staff, we expect that the findings for Cohort 3 depict a conservative representation of BARR implementation.

To meet school-level fidelity to the BARR model, schools were required to meet or exceed component thresholds for four of the five priority components that BARR Center deemed essential to BARR model implementation in the first year (i.e., restructuring, professional development, I-Time, block meeting, and risk review). Furthermore, schools were required to meet or exceed component thresholds for two of the three additional components that BARR Center deemed as supplementary to model implementation in the first year (i.e., family involvement, whole student emphasis, and contextual support). The expectation for cohort-level fidelity was that at least two thirds (67%) of schools met school-level fidelity. Appendix C provides information on component-, school-, and cohort-level thresholds as well as greater detail on how we calculated fidelity scores.

Across all three cohorts, 18 of 36[24] schools (50%) met or exceeded school-level fidelity during their first year of BARR implementation. Excluding four treatment schools that did not implement the BARR model during their assigned study year, 18 of 32 schools (56%) met school-level fidelity. This finding was below BARR Center's goal of at least 67% of schools achieving adequate implementation across the three cohorts. Most often, schools were successful at implementing the BARR components of risk review, professional development, and contextual support. Schools most often did not implement block meetings, family involvement, and whole student emphasis with fidelity. More detailed findings about implementation of the BARR model by cohort are presented in the following sections and in Appendix C.

Cohorts 1, 2, and 3 Implementation Fidelity. Exhibit 2.2 presents the number and percentage of schools by cohort that met component-level fidelity as well as school-level fidelity. Schools across the three cohorts were most often able to implement professional development. All three cohorts were least successful in implementing block meetings.

Cohort 1 Implementation Fidelity. In Cohort 1, 15 schools were randomly assigned to the treatment condition and were expected to implement the BARR model during the 2017–18 school year. However, three schools did not feel prepared to implement the model during the 2017–18 school year and elected to either forego implementation completely or postpone implementation to the 2018–19 school year. To capture these schools' decisions to not implement the BARR model, we assigned these schools a score of 0 for each of the eight components of BARR implementation.

[24] Thirty-seven schools were randomly assigned to the treatment condition and were expected to implement BARR during their initial study year. However, the fidelity of implementation analysis includes 36 schools because we were unable to secure a complete data set from one school in Cohort 3.

Exhibit 2.2. Number of Schools That Met School-Level Fidelity for Each BARR Component Measured by Cohort

	Cohort 1		Cohort 2		Cohort 3	All Cohorts	
	Number of Schools Meeting School-Level Fidelity						
	Including Crossover Schools	Excluding Crossover Schools	Including Crossover Schools	Excluding Crossover Schools	Excluding Crossover Schools	Including Crossover Schools	Excluding Crossover Schools
Component							
Restructuring Schedules	9 (60%)	9 (75%)	8 (73%)	8 (80%)	6 (60%)	23 (64%)	23 (72%)
Professional Development	9 (60%)	9 (75%)	9 (82%)	9 (90%)	9 (90%)	27 (75%)	27 (84%)
I-Time Curriculum	9 (60%)	9 (75%)	6 (55%)	6 (60%)	9 (90%)	24 (67%)	24 (75%)
Block Meetings	6 (40%)	6 (50%)	4 (36%)	4 (40%)	4 (40%)	14 (39%)	14 (44%)
Risk Reviews	9 (60%)	9 (75%)	8 (73%)	8 (80%)	9 (90%)	26 (72%)	26 (81%)
Family Involvement	6 (40%)	6 (50%)	6 (55%)	6 (60%)	5 (50%)	17 (47%)	17 (53%)
Whole Student Emphasis	9 (60%)	9 (75%)	5 (45%)	5 (50%)	7 (70%)	21 (58%)	21 (66%)
Contextual Support	8 (53%)	8 (67%)	9 (82%)	9 (90%)	9 (90%)	26 (72%)	26 (81%)
Number of Components							
4–5 Priority	8 (53%)	8 (67%)	5 (45%)	5 (50%)	6 (60%)	19 (53%)	19 (59%)
2–3 Additional	10 (67%)	10 (83%)	7 (64%)	7 (70%)	8 (80%)	25 (69%)	25 (78%)
School-Level Fidelity	8 (53%)	8 (67%)	5 (45%)	5 (50%)	5 (50%)	18 (50%)	18 (56%)
Sample Size (Number of Schools)	15	12	11	10	10	36	32

Note: Three treatment schools in Cohort 1 chose to delay implementation by 1 year and are considered crossover schools. In Cohort 2, one treatment school chose not to implement BARR. All treatment schools in Cohort 3 implemented BARR.

Eight of the 15 Cohort 1 treatment schools (53%) met school-level implementation expectations during the 2017–18 school year by meeting or exceeding thresholds on at least four of the five priority components and two of the three additional components. Two schools met the threshold for adequate implementation of BARR's additional components but did not meet the threshold for the priority components. Excluding the three schools that did not implement the BARR model during the 2017–18 school year, eight of 12 schools (67%) met school-level fidelity.

This meets BARR Center's goal for **cohort-level** fidelity (67% of schools with adequate school-level implementation). A detailed table of Cohort 1 fidelity findings is included in Appendix C.

Cohort 1 schools most often met implementation expectations for the following priority components: restructuring, professional development, I-Time, and risk review. Of the three additional components, Cohort 1 schools most often adequately implemented the whole student emphasis and contextual support components. The components that Cohort 1 schools did not often successfully implement were block meetings and family involvement.

Cohort 2 Implementation Fidelity. Eleven schools were randomly assigned to the treatment condition in Cohort 2 and were expected to implement the BARR model during the 2018–19 school year. One school elected to forego BARR implementation completely and was assigned a score of 0 for each of the eight components of BARR implementation in our fidelity analysis.

Five of the 11 Cohort 2 treatment schools (45%) met school-level fidelity during the 2018–19 school year. Excluding the one school that did not implement the BARR model, five of 10 schools (50%) met school-level fidelity. Two schools achieved adequate implementation for all eight components, and three schools implemented seven of the eight components with fidelity. Two additional schools met the threshold for adequate implementation of BARR's additional components but did not meet the threshold for the priority components. This represented a decline in implementation fidelity when compared to Cohort 1. A detailed table of the Cohort 2 fidelity findings is included in Appendix C.

Like Cohort 1, Cohort 2 schools most often met implementation expectations for the following priority components: restructuring, professional development, I-Time, and risk review. Of the three additional components, Cohort 2 schools most often adequately implemented the family involvement and contextual supports components. The components that Cohort 2 schools did not often implement successfully were block meetings and whole student emphasis.

Measuring Fidelity of Implementation for All Study Schools Versus Measuring Fidelity of Implementation for Implementing Study Schools

The sample for the fidelity of implementation analysis that we present in this chapter primarily includes 36 schools that were randomly assigned to the treatment condition and were expected to implement the BARR model during their initial study year. Although four schools decided to postpone or forego BARR implementation, we included those schools in our analysis in order to align with our intent-to-treat approach in the impact analysis. We conducted a separate fidelity of implementation analysis that only included the 32 schools that implemented the BARR model during their initial study year to examine implementation at schools that actually implemented the model. Throughout this section, we report any notable differences between the fidelity of implementation for all study schools versus schools that implemented the model.

Cohort 3 Implementation Fidelity. In Cohort 3, 11 schools were randomly assigned to the treatment condition and implemented the BARR model during the 2019–20 school year. Unlike the previous two cohorts, these schools experienced the unique challenge of moving their instruction and BARR implementation entirely online in the spring of 2020 because of the COVID-19 pandemic. As a result, we were unable to conduct our usual data collection activities in spring 2020. Instead, we measured fidelity of implementation by averaging two data sources: (1) ratings of the eight BARR components from interviews conducted with BARR coordinators in fall 2020 and (2) ratings of the eight BARR components determined by BARR coaches during their final coaching visit before COVID-19 school closures.[25]

Because of missing data for one school, only 10 of the 11 schools are included in the Cohort 3 fidelity analysis. Five of these schools (50%) met school-level fidelity during the 2019–20 school year. Of the five schools that met school-level fidelity, two schools adequately implemented all eight components. In addition, four Cohort 3 schools achieved adequate implementation for either BARR's priority components or additional components but did not achieve adequate implementation for the other set of components. A detailed table of the Cohort 3 fidelity findings is included in Appendix C.

Like cohorts 1 and 2, Cohort 3 schools most often met implementation expectations for the following priority components: restructuring, professional development, I-Time, and risk review. Of the three additional components, Cohort 3 schools were most often able to implement the whole student emphasis and contextual supports components. The components that Cohort 3 schools did not often implement successfully were block meetings and family involvement.

Summary of the Implementation Findings

The school staff we interviewed highlighted several facilitators and barriers related to implementation of each of the eight BARR components. For the professional development component, respondents most frequently highlighted the foundational training and BARR coaching visits as facilitators of professional development, noting that BARR coaching visits provided schools with actionable feedback informed by observations of meetings and I-Time lessons. Difficulties with scheduling was a common barrier noted in relation to successful implementation of the professional development component. For the restructuring component, although a few schools noted that they had trouble developing a schedule that would allow 80% of their students to have the same teachers within a cohort, in general, schools shared that restructuring had helped them to become more team oriented. As part of restructuring, schools also reported some variation in whether they conducted block meetings and risk review meetings weekly and whether a school counselor was able to consistently be part of BARR meetings.

[25] Final coaching site visits occurred between December 2019 and March 2020.

The family involvement component was generally discussed positively in BARR schools, with respondents sharing that the structure of the BARR model made it easier for them to communicate with families more frequently. This was largely due to the collaboration between BARR teachers and staff. Insufficient time to reach out to family members was a barrier mentioned by some respondents to successful family engagement. For the I-Time component, teachers from nearly all BARR schools reported conducting I-Time lessons where at least some students shared information about themselves, took risks, and engaged with teachers. Although many BARR schools described high levels of student engagement with I-Time lessons, respondents from some schools reported that some students did not participate in discussions, did not share, or felt the activities were not engaging. Scheduling due to issues such as assessments is another barrier to successful implementation of I-Time lessons that some BARR schools reported.

When discussing the block and risk review components, some teachers identified the following facilitators: the meeting structure, teacher collaboration around students, and the staff who participated in the meetings. BARR schools also reported challenges and barriers to implementing the block and/or risk review meetings with fidelity. Many BARR schools indicated that insufficient time was a challenge to implementing block and risk review meetings. Sometimes, there was not enough time to conduct the meeting; other times, there was not enough time to get through the whole student caseload for a meeting. For the whole student emphasis component, many BARR schools reported that teachers successfully broadened their perception of their role from purely providing academic instruction to supporting all facets of students, from academic to nonacademic strengths and needs. Some BARR teachers reported confidence and competence in working with the whole student, while others did not.

For the contextual support component, in general, school staff and leaders perceived there to be adequate levels of support for the BARR model at their schools. Respondents from nearly all BARR schools reported that school administrators attended BARR block and/or risk review meetings and/or met with the BARR coordinator regularly to discuss the model's implementation. Respondents from BARR schools also reported that their BARR coordinator supported implementation in multiple, significant ways, including organizing meetings, collecting data and supplies that teachers needed, and drawing support for BARR from the administration.

Our analysis of the fidelity of implementation data highlights that of the five priority components, schools had the most success implementing professional development and risk review meetings, and the least success implementing block meetings. Our interview data provided some additional context for understanding why this might be. For the professional development component, even though some schools reported challenges associated with scheduling, the foundational training and BARR coaching visits were reported frequently as

facilitators of implementing the BARR model. Schools may have prioritized the implementation of the indicators of this component as they found them valuable. By contrast, schools struggled most with implementing block meetings. Our interview data are less helpful in understanding why block meetings were more challenging to implement than risk review meetings because respondents did not distinguish between the two types of meetings in interviews, nor did we ask them to compare the meeting types. Respondents reported that insufficient time and the frequency of meetings were challenges to implementing the block and risk review meetings, which may partly explain the difficulty of implementing these meetings with fidelity, but it does not explain why schools were more successful with risk review meetings than block meetings. This difference in implementation success may reflect the greater perceived urgency of risk review and the inclusion of nonteaching staff in these meetings, such as the principal, assistant principal, or other student support staff. Those additional staff can help with problem-solving and implementation of interventions. Administrators also may have more flexibility to participate in meetings than teachers. Another possible explanation for why risk review was easier to implement may be that similar student support meetings were already commonplace in many schools, whereas block meetings, focusing on all students, were less common outside of BARR.

The extent to which schools can implement the eight BARR components with fidelity determines in part how much impact the BARR model can have on teacher and student experiences and students' academic performance. It is important to recognize that there were three schools in Cohort 1 and one school in Cohort 2 that postponed or decided not to implement the BARR model until the following year. Cohort 3 schools experienced school campus closures due to COVID-19 in the spring of their first year of implementation. Schools in all three cohorts experienced implementation challenges, and none of the cohorts of assigned treatment schools met the school-level fidelity threshold of 67% for adequate implementation. When we exclude the three Cohort 1 schools that did not implement the BARR model during the assigned year, Cohort 1 met cohort-level fidelity (67% of schools with adequate school-level implementation). Overall, 50% of assigned treatment schools and 56% of implementing treatment schools met adequate school-level implementation. Taken together, these implementation challenges could mean that, across cohorts, the contrast in teachers' and students' experiences with BARR between treatment and control schools is smaller than intended. The interpretation of the impact estimates presented in the next chapter must take this into account.

Chapter 3. The Impacts of BARR on Teachers and Students

This chapter describes the impact of BARR on academic and behavioral outcomes of students and on teacher attitudes and experiences after one year of BARR implementation. As introduced in Chapter 1, these impact analyses use data from administrative school and district records, a PSAT assessment administered in the fall of tenth grade, and surveys administered to teachers in the spring of the implementation year. For Cohort 3, these teacher surveys were delayed until the fall of the following year because of the COVID-19 pandemic and related school closures.[26]

We conducted the impact analyses presented in this chapter separately by cohort to account for the fact that the composition of the cohorts and their BARR program experiences differed. We also generated across-cohort estimates using meta-analysis techniques that treat each cohort as a separate study. Where data allow, our primary focus is on these summary estimates that aggregate results across all three cohorts. However, due to the COVID-19 pandemic, we were not able to administer the PSAT to most students in Cohort 3. We therefore only present impacts on PSAT outcomes for Cohort 1 and Cohort 2.

Our impact estimates are based on an "intent-to-treat" (ITT) approach in which schools are considered treatment or control schools entirely based on their random assignment status. Thus, schools are considered "BARR schools" if they were assigned to BARR, regardless of whether they implemented the BARR program during the study's implementation year or whether they extended BARR services to all their ninth-grade students and teachers. Four schools assigned to the BARR treatment group (three in Cohort 1 and one in Cohort 2) did not implement the BARR model during the study year, and eight other schools only delivered the BARR model to a subset of their students (five in Cohort 1, two in Cohort 2, and one in Cohort 3). To account for the fact that this lack of full implementation attenuates the study's impact estimates, we conducted supplemental analyses to obtain complier average causal estimates (CACE) of BARR's impacts on student outcomes.[27] We present these estimates alongside the ITT estimates in the tables in this chapter and highlight them as we discuss the findings. For details about the analytical and statistical methods we used to estimate program impacts, including

[26] As discussed in Chapter 1, we also conducted student surveys at the end of ninth grade (and in the fall of tenth grade for Cohort 3). Due to survey nonresponse and other methodological concerns, we decided not to include results from these surveys in this chapter. More details about these data and our methodological concerns are in Appendix F.

[27] We do not present CACE analyses for teacher outcomes because (a) we had no data on whether and which teachers were exposed to BARR in schools that only partially implemented the model, and (b) we had no teacher survey data from the one crossover school in Cohort 2. As a result, teacher-level CACE analyses were only possible for Cohort 1 for which they showed modestly larger impacts across all teacher-level outcomes.

details about the CACE analyses and the meta-analytical methods we used to combine estimates across cohorts, please refer to Appendix A.

BARR Improved Student Outcomes and Teacher Experiences

In summary, the impact analyses show that assignment to BARR improved student academic outcomes, including credit completion, course failure, GPA, and PSAT scores. BARR also reduced chronic absenteeism. The size and statistical significance of these impact estimates varied across cohorts and student groups, but most impact estimates were positive and in line with prior evidence on the effectiveness of BARR (Borman et al., 2021; Bos et al., 2019; Corsello & Sharma, 2015).

BARR also improved a range of teacher experiences and attitudes, with the strongest effects in the areas of teacher collaboration, data use, and teachers' views of their school's supports. Positive effects on teacher outcomes occurred in all three cohorts, and many persisted after the COVID-19 pandemic closed schools for part of the 2019–20 school year, as evidenced by the delayed Cohort 3 survey results.

As expected, most estimated impacts on student outcomes were larger when we statistically adjusted them for the fact that some BARR schools did not implement the program in the study year and that other BARR schools did not include all their ninth-grade students in the program.

Each subsequent section of this chapter presents a subset of these results, including how impacts varied across cohorts and student groups.

Did BARR Change the School Environment and Teacher Experiences?

An objective of the BARR model is to change how teachers view and interact with their students and with each other. By creating and promoting structures and activities to bring teachers together and to deepen teachers' relationships with their students, BARR aims to enhance teacher efficacy and student engagement.

As shown in Exhibit 3.1, we found that the estimated effect of BARR on teacher collaboration with and view of colleagues, as measured through teacher self-report on a survey, was large and statistically significant in all cohorts (ES ranging from 0.47 to 1.01). Impacts on teacher data use were also substantial in all cohorts (ES ranging from 0.31 to 0.58). The impact on teachers' views of school's supports was statistically significant and substantial for Cohorts 2 and 3 (ES = 0.35 and 0.57, respectively), resulting in a positive and statistically significant impact on this outcome across the three cohorts (ES = 0.41). In addition, we found smaller, but statistically significant, impacts in the three cohorts combined on teachers' views of their students'

postsecondary expectations (ES = 0.25), students' actual behavior (ES = 0.23), perceptions of students' behavior (ES = 0.22), and having a more strength-based mindset (ES = 0.21).

Exhibit 3.1. Impacts on Teacher Experiences Measured by Teacher Surveys, by Cohort

	N	BARR	Control	Difference	p-value	Effect size
Cohort 1						
Student academic motivation	146	49.3	47.1	2.12	0.301	0.22
Teacher collaboration with and view of colleagues	143	55.1	45.0	10.16 ***	0.000	1.01
Teacher use of data	144	53.6	50.2	3.40 *	0.021	0.33
Postsecondary educational expectations	146	50.4	45.6	4.80 **	0.001	0.58
View of the school's supports	143	51.3	50.0	1.29	0.585	0.13
Teacher self-efficacy	144	53.8	52.8	0.97	0.393	0.11
Strength-based mindset	147	51.6	49.0	2.53 *	0.043	0.28
Students' actual behavior	145	48.9	47.0	1.89	0.211	0.20
Perception of students' behavior	146	48.6	46.3	2.33	0.114	0.25
Cohort 2						
Student academic motivation	202	49.6	47.2	2.37	0.337	0.24
Teacher collaboration with and view of colleagues	199	50.6	46.0	4.51 **	0.001	0.47
Teacher use of data	200	50.0	47.0	3.05 *	0.026	0.31
Postsecondary educational expectations	205	49.5	48.6	0.91	0.730	0.08
View of the school's supports	200	49.7	46.1	3.59 *	0.026	0.35
Teacher self-efficacy	199	51.4	52.2	-0.83	0.319	-0.09
Strength-based mindset	206	50.9	49.2	1.75	0.095	0.18
Students' actual behavior	202	49.8	47.3	2.44 **	0.006	0.24
Perception of students' behavior	202	50.7	46.5	4.19 *	0.012	0.43
Cohort 3						
Student academic motivation	158	52.9	54.6	-1.71	0.307	-0.17
Teacher collaboration with and view of colleagues	153	54.4	47.7	6.75 **	0.006	0.75
Teacher use of data	156	52.2	46.9	5.30 *	0.010	0.58
Postsecondary educational expectations	159	52.9	52.7	0.13	0.927	0.01
View of the school's supports	152	54.1	49.0	5.10 ***	0.000	0.57
Teacher self-efficacy	153	44.5	44.5	-0.01	0.988	0.00
Strength-based mindset	159	50.3	48.5	1.73	0.412	0.15
Students' actual behavior	156	54.6	52.3	2.39	0.157	0.25
Perception of students' behavior	157	54.2	53.7	0.47	0.730	0.05

(continues)

Exhibit 3.1. Impacts on Teacher Experiences Measured by Teacher Surveys, by Cohort (continued)

	N	BARR	Control	Difference	p-value	Effect size
Cohorts 1, 2, and 3						
Student academic motivation	506	51.0	50.7	0.37	0.746	0.04
Teacher collaboration with and view of colleagues	495	52.1	46.2	5.95 ***	0.000	0.63
Teacher use of data	500	51.8	48.2	3.62 ***	0.000	0.36
Postsecondary educational expectations	510	51.4	49.0	2.32 *	0.014	0.25
View of the school's supports	495	52.1	48.2	3.93 ***	0.000	0.41
Teacher self-efficacy	496	49.5	49.6	-0.13	0.808	-0.01
Strength-based mindset	512	51.1	49.0	2.03 **	0.006	0.21
Students' actual behavior	503	50.4	48.1	2.31 **	0.001	0.23
Perception of students' behavior	505	51.4	49.3	2.10 *	0.014	0.22

Source: AIR calculations from AIR-administered teacher surveys. Scales are standardized to a mean of 50 and standard deviation of 10.

Note. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. Combined Cohorts 1, 2, and 3 results were obtained using a meta-analysis model that averaged the estimated difference across cohorts giving more weight to estimates with greater precision. For additional details, see Appendix D.

Did BARR Change Outcomes for Students?

Impacts on Credit Attainment, Course Failure, and Grades

By changing the school and classroom environment and increasing the academic and nonacademic supports available to students, BARR aims to enhance students' motivation and classroom engagement and improve academic performance. This is expected to manifest in fewer failed courses, more credits earned, and a higher GPA, as well as better academic achievement.

To determine the impact of BARR on these outcomes, we collected transcript data for all ninth-grade students in the study schools. We used these data to create three measures for each student: (1) the percentage of core credits earned, (2) a 0/1 indicator of whether students passed all their core courses, and (3) their GPA. We present full sample estimates (by cohort and for the three cohorts combined) for each of these three measures in Exhibit 3.2 and then present estimates for different groups of students for each outcome across the combined cohorts in Exhibit 3.3. Detailed subgroup estimates by cohort are presented in Appendix E.

As shown in the table's bottom panel, which summarizes impact estimates across the three cohorts, students in BARR schools earned a greater percentage of credits for the core courses

they enrolled in: 89.1 percent compared to 85.5 percent for students in control schools (ES = 0.14). Students in BARR schools also were more likely to have passed all their courses: 80.2 percent compared to 74.4 percent (ES = 0.21). BARR students also had a modestly higher average ninth-grade GPA, 2.6 compared to 2.5 for control students (ES = 0.13). We found that the positive impacts on these three outcomes were mostly limited to Cohort 2 and Cohort 3. In Cohort 1, estimated effects for these outcomes were small and not statistically significant. The positive impact on GPA was mostly accounted for by Cohort 3.

Exhibit 3.2. Estimated Impact of BARR on Credit Attainment, Course Failure, and GPA, by Cohort

Outcome	N	BARR	Control	Difference	p-value	Effect size	*Compliance-adjusted*	
							Difference	*Effect size*
Cohort 1								
Core credits earned (%)	6,003	86.7	84.8	1.90	0.508	0.07	3.32	0.13
Passed all courses (%)	6,003	73.4	71.3	2.09	0.741	0.06	3.49	0.11
GPA	5,915	2.5	2.6	-0.10	0.450	-0.10	-0.19	-0.18
Cohort 2								
Core credits earned (%)	6,822	84.6	79.4	5.22*	0.013	0.18	7.66**	0.27
Passed all courses (%)	6,822	67.3	58.9	8.43*	0.034	0.22	11.83*	0.32
GPA	6,859	2.4	2.3	0.08	0.321	0.07	0.12	0.12
Cohort 3								
Core credits earned (%)	6,703	91.0	87.7	3.39**	0.005	0.14	3.80**	0.16
Passed all courses (%)	6,703	87.8	82.8	4.98**	0.001	0.24	5.27**	0.27
GPA	6,842	2.8	2.6	0.23***	0.000	0.22	0.25***	0.24
Cohorts 1, 2, and 3 Combined								
Core credits earned (%)	19,528	89.1	85.5	3.62***	0.000	0.14	4.41***	0.18
Passed all courses (%)	19,528	80.2	74.4	5.78***	0.000	0.21	6.63***	0.28
GPA	19,616	2.6	2.5	0.13**	0.004	0.13	0.19**	0.18

Source: AIR calculations from school-provided administrative data.

Note. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. Combined Cohorts 1, 2, and 3 results were obtained using a meta-analysis model that averaged the estimated difference across cohorts giving more weight to estimates with greater precision. For additional details, see Appendix E.

When we calculated *CACE* estimates to account for the fact that some schools and students assigned to the BARR group were not exposed to the program as intended, we found that impacts on these three outcomes became more favorable, with the effect size on passing all courses increasing to 0.28 for the full sample across the three cohorts and the effect sizes for the other two outcomes (percentage of credits earned and GPA) increasing to 0.18.

The favorable estimates for the impacts on credit attainment and passing all courses were consistent across many different student groups and statistically significant for most (Exhibit 3.3).

Exhibit 3.3. Estimated Impact of BARR on Credit Attainment, Course Failure and GPA, by Student Group – Cohorts 1, 2, and 3 Combined

Student Group	N	BARR	Control	Difference	p-value	Effect size	*Compliance-adjusted*	
							Difference	*Effect size*
Core Credits Earned (%)								
Female	9,380	91.7	88.3	3.4***	0.001	0.15	4.3***	0.20
Male	10,027	86.7	82.8	3.9***	0.000	0.14	4.5***	0.16
Students of color	12,701	87.1	83.1	4.0***	0.000	0.14	4.8***	0.17
White	6,537	94.4	91.9	2.5*	0.024	0.13	3.7**	0.19
Free or reduced-price lunch	9,996	86.1	80.3	5.7***	0.000	0.21	6.5***	0.23
English learners	2,909	79.2	73.7	5.6***	0.000	0.18	2.7	0.08
Special education	2,804	83.7	78.3	5.5***	0.000	0.18	5.1*	0.17
Passed All Courses (%)								
Female	9,380	84.6	80.0	4.6***	0.000	0.20	5.3***	0.27
Male	10,027	75.5	68.3	7.2***	0.000	0.23	8.4***	0.30
Students of color	12,701	77.5	70.8	6.8***	0.000	0.23	8.0***	0.29
White	6,537	88.2	85.6	2.6	0.153	0.14	3.0	0.19
Free or reduced-price lunch	9,996	69.0	56.9	12.1***	0.000	0.33	13.8***	0.40
English learners	2,909	62.8	54.0	8.8**	0.003	0.22	7.2*	0.18
Special education	2,804	67.5	59.0	8.4**	0.008	0.23	9.0*	0.25
GPA								
Female	9,406	2.82	2.68	0.14**	0.002	0.14	0.20***	0.20
Male	10,088	2.47	2.34	0.13**	0.007	0.12	0.19**	0.18
Students of color	12,922	2.56	2.33	0.24***	0.000	0.22	0.33***	0.30
White	6,120	2.96	2.96	0.00	0.973	0.00	0.00	0.00
Free or reduced-price lunch	9,987	2.45	2.24	0.21***	0.000	0.21	0.25***	0.24
English learners	2,985	2.19	1.92	0.27***	0.000	0.28	0.24***	0.25
Special education	2,811	2.14	2.08	0.07	0.219	0.06	0.13	0.12

Source: AIR calculations from school-provided administrative data.

Note. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. Combined Cohorts 1, 2, and 3 results were obtained using a meta-analysis model that averaged the estimated difference across cohorts giving more weight to estimates with greater precision. For additional details, see Appendix E.

The impact of BARR on the percentage of credits earned was greatest for students eligible for free or reduced-price lunch, for whom this percentage was 86.1 in the BARR schools compared to 80.3 in the control schools (ES = 0.21). Similarly, 69.0 percent of BARR students eligible for

free or reduced-price lunch passed all their core courses, compared to 56.9 percent of their counterparts in the control group (ES = 0.33). Impacts on GPA were strongest for English learners (ES = 0.28). Across all three outcomes, estimated impacts were smallest for white students (ES = 0.13 and 0.14 for credits earned and passing all core courses, and no impact on GPA for these students). Outcomes for white students in the control group left less room for improvement—control students earned 91.9 percent of their credits and 85.6 percent of them passed all their courses. White students also had the highest GPA of any of the student groups (2.96 for BARR students and control students).

One way to look at these patterns of impacts across student groups is to examine the extent to which BARR reduced disparities in these academic outcomes between different groups of students. For example, in control schools, students eligible for free or reduced-price lunch were more than twice as likely to have a course failure than students not eligible for free or reduced-price lunch (43.1 versus 17.4 percent in control schools[28]). BARR significantly reduced this gap. In BARR schools, only 31.0 percent of students eligible for free or reduced-price lunch failed a course, compared to 12.3 percent of students not eligible for free or reduced-price lunch. Thus, BARR reduced the gap in this outcome between these two student groups from 25.6 to 18.7 percentage points, a 27 percent reduction.

Impacts on PSAT Reading and Mathematics Scores

To help us determine the impact of BARR on students' academic achievement, schools agreed to administer the PSAT to all tenth-grade students in the study schools in the fall of their 10th-grade year. Unfortunately, the COVID-19 pandemic seriously compromised PSAT administration in fall 2020. Most Cohort 3 schools did not administer the PSAT that fall, and in most schools that did administer the assessment, only a small subset of students participated in it. Because of this, we did not conduct PSAT impact analyses for Cohort 3 and could only estimate cross-cohort impacts for Cohorts 1 and 2 together. This resulted in reduced statistical power for these PSAT impact analyses, which means that we were unable to detect small impacts that may have been statistically significant absent these pandemic-related challenges.

In Exhibit 3.4, we present impacts on PSAT reading, math, and total scores for Cohort 1 and Cohort 2.[29] We found that the PSAT impacts were favorable for Cohort 1 (ES = 0.14 for the

[28] We calculate these rates by subtracting the control group proportions of "passed all courses" from 100 for these student groups. To see estimated outcomes for students not eligible for free or reduced-price lunch, please refer to the complete impact results in Appendix E.

[29] As noted in Chapter 1, we only included schools that administered the PSAT-NMSQT test in the fall of 10th grade in these analyses. Three schools in Cohort 1 and one school in Cohort 2 administered the PSAT-10 (a different test) in the spring of 10th grade. To minimize the complexity of the impact analysis, we dropped these four schools, but we conducted a sensitivity analysis to ensure that doing so would not materially change the impact estimates. The results from this analysis are in Appendix E.

mathematics subtest and for the total PSAT) but small and not statistically significant for Cohort 2. The impact on the total PSAT for the two cohorts combined (ES = 0.07) was not statistically significant (p-value 0.057), but there was a positive and statistically significant impact on the PSAT math score (ES = 0.07; p-value 0.016). The pattern of impacts across cohorts differs from the earlier impact estimates for credit attainment and course failure where BARR had stronger impacts in Cohort 2. Adjusting for noncompliance at the school and student levels increased the estimated effect sizes for Cohort 1 (to ES = 0.17 for the math subtest and the full PSAT score).

Exhibit 3.4. Estimated Impact of BARR on PSAT Scores, by Cohort

Outcome	N	BARR	Control	Difference	p-value	Effect size	*Compliance-adjusted*	
							Difference	*Effect size*
Cohort 1								
Reading	3,443	433	422	11.0	0.055	0.13	13.4	0.15
Mathematics	3,443	421	410	10.7 **	0.001	0.14	13.1 **	0.17
Total	3,443	854	832	21.6 *	0.014	0.14	26.4 *	0.17
Cohort 2								
Reading	3,829	425	422	2.3	0.631	0.03	3.1	0.04
Mathematics	3,829	424	423	1.0	0.761	0.01	1.3	0.02
Total	3,829	849	845	3.3	0.650	0.02	4.5	0.03
Cohort 1 and Cohort 2								
Reading	7,272	428	422	5.9	0.109	0.07	7.6	0.09
Mathematics	7,272	422	417	5.5 *	0.016	0.07	7.1 *	0.09
Total	7,272	851	840	10.6	0.057	0.07	13.5	0.09

Source: AIR calculations from school-provided PSAT data.

Note. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. Combined Cohorts 1 and 2 results were obtained using a meta-analysis model that averaged the estimated difference across cohorts giving more weight to estimates with greater precision. For additional details, see Appendix E.

In Exhibit 3.5, we present impacts on the PSAT total score for different student groups for the two cohorts combined. We found statistically significant positive effects on total PSAT scores for male students (ES = 0.11), students of color (ES = 0.13), students eligible for free or reduced-price lunch (ES = 0.12), and English learners (ES = 0.19).

Exhibit 3.5. Estimated Impact of BARR on PSAT Total Scores, by Student Group – Cohorts 1 and 2 Combined

Student Group	N	BARR	Control	Difference	p-value	Effect size	*Compliance-adjusted*	
							Difference	*Effect size*
Female	3,671	854	852	1.9	0.718	0.01	3.4	0.02
Male	3,601	846	830	16.6 *	0.020	0.11	20.6 *	0.13
Students of color	4,362	819	802	17.0 **	0.002	0.13	21.0 **	0.16
White	2,902	894	902	-8.2	0.061	-0.05	-1.5	-0.01
Free or reduced-price lunch	3,435	814	799	14.4 ***	0.001	0.12	17.5 ***	0.14
English learners	1,091	767	749	18.1 **	0.008	0.19	20.4 *	0.21
Special education	733	735	738	-2.7	0.701	-0.02	-3.1	-0.03

Source: AIR calculations from school-provided PSAT data.

Note. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. Combined Cohorts 1 and 2 results were obtained using a meta-analysis model that averaged the estimated difference across cohorts giving more weight to estimates with greater precision. For additional details, see Appendix E.

Impacts on Attendance, Suspensions, and Persistence to 10th Grade

The last set of outcomes we examined included chronic absenteeism, suspensions, and persistence to 10th grade. We used administrative data from schools and districts to create these outcome measures. As described in Chapter 1, we restricted the attendance and suspension data for Cohort 3 to the proportion of the year preceding the pandemic-related campus closures, where possible. To capture persistence to 10th grade, we measured whether students enrolled in 10th grade in the same school. Some students may have transferred to another school for 10th grade, so this is a conservative measure of persistence to 10th grade. We did not include impact estimates for persistence to 10th grade for Cohort 3 because these data were unreliable due to the 2020–21 school year starting virtually for most students.

In Exhibit 3.6 we show impacts of BARR on chronic absenteeism, suspension, and persistence to 10th grade for the three individual cohorts and for the three cohorts combined. We found the strongest effects on these outcomes in Cohort 1, where students in BARR schools had significantly lower suspensions. Of students in control schools, 11.6 percent were suspended, compared to 7.6 percent of students in BARR schools (ES = –0.29). We also found similarly large favorable effects on absenteeism and persistence to 10th grade in Cohort 1 (ES = –0.26 and 0.35, respectively), but those latter two effects were not statistically significant (p-values of 0.072 and 0.090, respectively). All three of these estimates were considerably more favorable once adjusted for the fact that many students in schools assigned to BARR in Cohort 1 did not have access to the program. In Cohort 2, we found no effects on any of these three outcomes. In Cohort 3, we found that students in BARR schools had significantly lower chronic

absenteeism than students in control schools. Approximately one in seven BARR students in Cohort 3 was chronically absent (13.2 percent), compared to approximately one in five of their counterparts in control schools (19.0 percent), for an effect size of –0.27. There were no impacts on suspensions in Cohort 3. Across all three cohorts combined, this pattern of impacts translates into lower chronic absenteeism for students in BARR schools (ES = –0.16).

Exhibit 3.6. Estimated Impact of BARR on Behavioral Outcomes, by Cohort

Outcome	N	BARR	Control	Difference	p-value	Effect size	*Compliance-adjusted*	
							Difference	*Effect size*
Cohort 1								
Chronically absent (%)	6,303	12.6	18.0	-5.4	0.072	-0.26	-8.4	-0.46
Ever suspended (%)	5,841	7.6	11.6	-4.0 *	0.031	-0.29	-5.3 *	-0.46
Persistence to 10th grade (%)	6,305	94.5	90.7	3.8	0.090	0.35	5.4	0.62
Cohort 2								
Chronically absent (%)	7,533	24.1	24.6	-0.4	0.866	-0.01	-0.6	-0.02
Ever suspended (%)	5,671	9.3	8.4	0.9	0.568	0.07	1.2	0.10
Persistence to 10th grade (%)	7,555	85.9	82.7	3.2	0.268	0.14	4.7	0.23
Cohort 3								
Chronically absent (%)	7,519	13.2	19.0	-5.9 *	0.028	-0.27	-6.5 *	-0.30
Ever suspended (%)	7,392	4.2	3.3	1.0	0.241	0.16	1.1	0.18
Cohorts 1, 2, and 3 Combined								
Chronically absent (%)	21,355	19.0	21.8	-2.8 *	0.041	-0.11	-3.9 *	-0.16
Ever suspended (%)	18,904	7.3	8.0	-0.7	0.818	-0.06	-0.3	-0.03
Cohorts 1 and 2 Combined								
Persistence to 10th grade (%)	13,860	88.4	85.0	3.3	0.065	0.18	4.9	0.28

Source: AIR calculations from school-provided administrative data.
Note. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. Combined Cohorts 1, 2, and 3 and Combined Cohorts 1 and 2 results were obtained using a meta-analysis model that averaged the estimated difference across cohorts giving more weight to estimates with greater precision. For additional details, see Appendix E.

In Exhibit 3.7, we show how BARR's impacts on these three outcomes varied across student groups. We found that BARR significantly reduced absenteeism for male students, students of color, and English learners (reductions of 3.4, 3.6, and 2.7 percent in the percentage of students chronically absent for these three groups, respectively). This translates into effect sizes of –0.13 for both male students and students of color and –0.09 for English learners , or –0.18 for male students, –0.16 for students of color, and –0.05 for English learners after adjusting for noncompliance with assignment status. The lower rates of chronic absenteeism for students of color in BARR schools (21.1 percent compared to 24.7 percent in control schools) translated into a substantially smaller gap in chronic absenteeism between students of color and white students in BARR schools, which we estimated to be a gap of 8.2 percentage points, compared

to control schools for which the equivalent gap was 12.8 percentage points, a 36 percent reduction.

Exhibit 3.7. Estimated Impact of BARR on Behavioral Outcomes, by Student Group – Combined Cohorts

Student Group	N	BARR	Control	Difference	p-value	Effect size	*Compliance-adjusted*	
							Difference	*Effect size*
Chronically absent (%)								
Female	10,174	18.1	20.8	-2.7	0.089	-0.11	-3.6	-0.15
Male	10,968	18.7	22.0	-3.4 *	0.016	-0.13	-4.7 *	-0.18
Students of color	14,134	21.1	24.7	-3.6 *	0.033	-0.13	-4.5 *	-0.16
White	6,827	12.9	11.9	1.0	0.235	0.06	1.2	0.07
Free or reduced-price lunch	10,920	23.0	25.7	-2.7	0.111	-0.09	-3.8	-0.13
English learners	3,284	22.7	25.4	-2.7 *	0.014	-0.09	-1.6	-0.05
Special education	3,129	32.3	33.5	-1.1	0.607	-0.03	-1.6	-0.05
Ever suspended (%)								
Female	9,039	4.8	4.7	0.2	0.441	0.02	0.4	0.06
Male	9,652	9.2	11.2	-2.0	0.144	-0.13	-1.6	-0.13
Students of color	11,703	8.0	8.2	-0.2	0.841	-0.02	0.2	0.02
White	6,811	5.2	6.3	-1.2	0.861	-0.14	-1.2	-0.14
Free or reduced-price lunch	9,106	9.7	9.7	0.1	0.448	0.00	0.5	0.04
English learners	2,662	8.3	9.5	-1.1 *	0.030	-0.09	-1.2	-0.12
Special education	2,781	15.8	13.5	2.4	0.060	0.12	4.8 *	0.25
Persistence to 10th grade in same school (%) †								
Female	6,563	89.1	86.2	2.9	0.066	0.17	4.5	0.27
Male	7,097	88.2	84.5	3.7	0.076	0.19	5.1	0.28
Students of color	9,433	85.5	81.2	4.3	0.077	0.19	6.1	0.27
White	4,052	95.7	94.1	1.6	0.361	0.21	1.6	0.27
Free or reduced-price lunch	7,380	88.4	82.9	5.5 **	0.008	0.28	6.7 **	0.40
English learners	2,382	80.4	77.7	2.7	0.478	0.10	3.2	0.12
Special education	2,063	86.2	82.9	3.3	0.055	0.16	5.1	0.26

Source: AIR calculations from school-provided administrative data.

Note. † = Results provided for Cohorts 1 and 2 only. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. For additional details, see Appendix E.

Looking at suspensions, we found that BARR significantly reduced suspensions for English learners (reduction of 1.1 percent in the percentage of students suspended for this group). Suspension rates were lower in BARR schools than in control schools for most student groups, except for

students in special education, for whom we observed an unfavorable impact that became statistically significant once we adjusted for noncompliance with assignment to BARR (ES = 0.25).

Lastly, we found statistically significant positive effects on persistence to 10th grade for students eligible for free or reduced-price lunch (ES = 0.28) for combined Cohorts 1 and 2. This translates into an effect size of 0.40 for students eligible for free or reduced-price lunch after adjusting for noncompliance with assignment status.

Summary

Schools assigned to the BARR model did better than schools assigned to the control group on a range of important outcomes for students and teachers. Students performed better academically, earning higher grades, passing more courses, and demonstrating stronger academic achievement. They also were less likely to be chronically absent. Teachers had more favorable views of their colleagues and students and were more likely to collaborate with their colleagues. They were also more likely to use data to guide their instruction and their work with students, and they were more positive about the school supports available to them.

These favorable program effects were usually moderate to large (effect sizes ranging from 0.2 to 0.5) and were evident across multiple cohorts and student groups and stronger when adjusted for noncompliance with assignment to BARR. However, there were exceptions; the size and statistical significance of our impact estimates varied across cohorts and student groups and some student groups benefited more from their exposure to BARR than others. BARR's program effects were generally stronger for students of color and students eligible for free or reduced-price lunch and were weaker for white students. Male students also benefited more than female students, although those differences were generally small.

In Chapter 5, we discuss some of the caveats associated with this impact analysis and summarize our findings and their implications across the evaluation as a whole.

Chapter 4. Scaling Up the BARR Model

As stated in Chapter 1, we documented six strategies BARR Center planned to use to grow the model over time and sustain long-term positive outcomes for schools that participate in the model. The two strategies for growing the program to more schools included fortifying BARR Center infrastructure and expanding school and district awareness of the model. The four strategies used to ensure high-quality implementation as the model grew included sustaining cost of service, training of coaches at scale, school-level staff training at scale, and school-level BARR implementation over time.[30]

We met annually with BARR Center staff to understand and document plans to grow and sustain the model during and after the grant award period. These conversations were the basis for the scale-up measurement framework. The design of the framework was developed by first identifying the most critical barriers to scaling noted by BARR Center. These barriers to scaling were then translated into strategies to grow the model over time and sustain long-term positive BARR model outcomes for schools that use the model. Appendix G provides additional details on the indicators measured under each of the six scale-up strategies.

To understand the extent to which BARR scale-up strategies were completed as planned during the grant award period, we requested communication records, school and coach training attendance records, staffing information, and BARR Center-collected assessments of school implementation of the BARR model from BARR Center and the Hazelden Betty Ford Foundation.[31]

What efforts did BARR Center make to scale up the BARR model?

There were two main strategies to grow the program: fortifying BARR Center infrastructure and expanding school and district awareness of the program. To address the first strategy, **fortifying BARR Center infrastructure,** BARR Center believed they needed to maintain adequate staffing, operate on a fiscally sustainable budget, and build BARR model visibility at the national level. In addition, BARR Center believed it was critical to engage with schools and districts that were already implementing the program to **expand school and district awareness of the model** to more schools**.** This strategy was intentionally measured separately from efforts that BARR Center engaged in to promote program visibility because success was reliant on schools that had implemented the program for at least a year and was predicated on community

[30] During the first two years of the grant, BARR Center further refined the strategies of scaling and the processes for measuring their progress on addressing those needs. This report presents our findings on the six strategies that BARR Center ultimately deemed essential for BARR scale-up.

[31] BARR Center partners with the Hazelden Betty Ford Foundation to hire and train BARR coaches. These coaches work directly with BARR coordinators employed by the schools and/or districts implementing the BARR model.

relationships among local schools and districts. To encourage schools and districts to support efforts to grow awareness of the BARR model, BARR Center planned to share effective communication strategies with schools and encouraged study schools to use this guidance to promote their experience with the model.

The four strategies related to sustaining the quality of model implementation as BARR grows include sustainable cost of services, school-level staff training, coach training, and school implementation of the model. To address **cost of service**, BARR Center worked with the Hazelden Betty Ford Foundation to determine the cost associated with maintaining the program after grant funding ended. These maintenance costs give participating schools access to the BARR resource portal (including 100+ I-Time lessons), virtual and on-site coaching, professional learning communities (PLCs) and video training, and a reduced rate to attend the BARR National Conference to encourage continued high-quality implementation of the program. Indicators for this strategy included sharing of detailed guidance about the cost of services with study schools and actively working with schools interested in continuing the model after the grant funding ended to secure additional funding to pay for the model. To ensure **training of coaches at scale**, BARR Center aimed to use a mentoring model to prepare their growing number of coaching staff to work directly with school-based BARR coordinators and other school staff at any school that adopted the model. BARR Center planned to create a rubric as a framework for examining coaches' proficiency in essential knowledge areas and to standardize coaching expectations and practices. Similarly, to ensure **school-level staff training at scale** (i.e., ensuring that school staff have the skills and knowledge to implement the model), BARR Center planned to work with coaches to ensure that minimum training requirements were completed annually by schools implementing the model. To accomplish the final strategy, which is sustainable **school-level BARR implementation** within schools, BARR Center planned to monitor and support implementation annually and to formalize a school accreditation process for schools that met or exceeded BARR Center's standards for adequate implementation.

To measure the extent to which BARR scale-up strategies were completed as planned, we used program data provided by BARR Center and Hazelden Betty Ford Foundation to rate progress toward reaching each scale-up strategy goal.[32] We assessed the scale-up of the BARR model in 4 of the 5 years of the grant period (June 2017–May 2021).[33] Measurement was based on

[32] The data sources for scale-up measurement included interviews with BARR regional managers and coaches as well as BARR data, such as communication records, training attendance, staffing information, and implementation scores collected by BARR coaches. As a result, some findings presented in this section may differ slightly from analyses presented elsewhere in this report, which were conducted with data collected solely by AIR researchers.

[33] BARR Center operates and provides services to schools outside of a school-year calendar. To ensure that we captured all efforts to grow and sustain the BARR model, we measured scale-up strategies with a timeline of calendar years rather than cohorts of schools.

predetermined thresholds set for each indicator of BARR scale-up. All thresholds for this analysis were established in consultation with BARR Center and based on theoretical expectations and practical experiences before and during the study's first two years. Appendix G provides more detail on data sources, indicators, and thresholds for scale-up measurement.

Exhibit 4.1 summarizes the extent to which BARR Center succeeded in growing and sustaining the model during the grant period. BARR Center demonstrated progress on all six strategies introduced above during the grant period. BARR Center was particularly successful in fortifying their infrastructure, establishing and sharing the cost of services, and scaling school-level and coach training and supports. BARR strategies related to expanding school-level awareness of BARR and school-level implementation, the two strategies that were primarily driven by school administrator decisions and not within the direct control of BARR Center, were less successfully scaled. These two strategies were primarily driven by school administrator decisions and not within the direct control of the BARR Center. As such, BARR Center could only provide limited support to scaling these strategies.

Exhibit 4.1. Progress Towards Growing and Sustaining the BARR Model by Strategy and Implementation Year

Strategies and Indicators	Year 1 (2017–18)	Year 2 (2018–19)	Year 3 (2019–20)	Year 4 (2020–21)
Strategies for Expanding the BARR Model				
Strategy: Fortify BARR Center Infrastructure				
a. BARR Center maintains staff-to-school ratio	n/a	●	○	●
b. BARR Center fills vacancies and open positions	n/a	●	●	●
c. BARR Center is fiscally sustainable	n/a	●	●	●
d. BARR Center builds BARR model visibility	n/a	●	●	●
Strategy: Expand School and District Awareness				
a. BARR Center shares school communication strategies	*	*	*	*
b. Schools use BARR communication strategies	n/a	◘	◘	◘
Strategy: Sustain Cost of Services				
a. BARR Center shares cost of services/ materials	n/a	n/a	●	◘
b. Schools in need of supplemental funding secure funds for the subsequent year	n/a	n/a	●	●

(continues)

Exhibit 4.1. Progress Towards Growing and Sustaining the BARR Model by Strategy and Implementation Year *(continued)*

Strategies and Indicators	Year 1 (2017–18)	Year 2 (2018–19)	Year 3 (2019–20)	Year 4 (2020–21)
Strategy: Sustain Coach Training at Scale				
a. New and continuing coaches complete BARR training	●	●	●	●
b. BARR Center creates coach mastery rubric	n/a	●	●	●
Strategy: Sustain School-Level Staff Training at Scale				
a. Educators complete BARR training	●	●	●	●
b. Schools use BARR database	n/a	◐	◐	◐
Strategy: Sustain School-Level Implementation				
a. Schools implement BARR model with fidelity	◐	◐	◐	◐
b. BARR Center creates school accreditation process	n/a	n/a	●	●

Note: = school year; ○ = Did not meet, ◐ = Partially met, ● = Met.
*BARR Center held sessions on communication strategies with schools during their annual conference. However, BARR Center did not document attendance at these sessions, so we were not able to confirm whether this indicator was completed as planned.

More detailed findings about BARR Center's progress on each of these scale-up strategies are presented in the following sections.

Strategies for Expanding the BARR Model

Strategy: Fortify BARR Center Infrastructure

As the BARR model expands to more schools and districts, it is critical that BARR Center has a sustainable plan for addressing staffing, ensuring fiscal health, and promoting awareness of the BARR model. BARR Center identified specific staff-to-school coaching ratios and the ability to fill BARR Center staffing vacancies as important factors for scaling the model. The measure used to examine fiscal health includes BARR Center's ability to secure 501(c)3 status and remain fiscally sound as demonstrated by annual financial statements (990 series). Finally, BARR Center expected that successful scale-up would require them to increase BARR model visibility, in addition to building BARR model awareness via schools and districts, through presentations,

local news articles, social media, website updates, and partnerships with national organizations. For each year of scale-up measurement, BARR Center met or exceeded these expectations.[34]

When asked about resources that would be needed by BARR Center to continue scaling up the BARR model, nearly all interviewed regional managers and coaches said that more staff (particularly coaches) were needed. Interviewees also expressed a need for more support in expanding awareness of the BARR model. They noted that BARR Center could use help to increase their social media presence, raise conference attendance, and build connections to local news outlets. As more schools implement BARR, regional managers and coaches in Year 2 and Year 3 interviews also suggested organizing schools and coaching assignments into regional networks. This would provide coaches with easier access to nearby schools and help facilitate connections among schools in the area. Some respondents also suggested ways to support implementation, such as an online database of BARR resources that schools and coaches could access at any time.

Communication Strategies Used by BARR Center and Schools to Expand Awareness of the Model

BARR Center developed and used several communication strategies to promote the BARR model. These included the following:

- **BARR website**: Each BARR school has a page on the website that includes basic school information, BARR implementation details, and testimonials from school staff.
- **Monthly newsletter**: BARR Center shares updates on their recent activities and success stories from schools that use the model. The mailing list includes over 4,000 school staff, families, school support organizations, funders, and interested schools.
- **Professional learning communities:** BARR Center provides professional development and networking opportunities to BARR school staff and, on certain occasions, the general public.
- **Thought leadership opportunities**: BARR Center offers learning opportunities through presentations, op-eds, podcasts, and videos that are shared with their larger network.

Strategy: Expand School and District Awareness

BARR Center, in part, builds school and district awareness and expansion of the model by encouraging current BARR schools to communicate their BARR implementation and impact stories to other schools and districts in their communities. BARR Center believes it is critical (1) to share and promote ways in which schools can build greater awareness[35] and (2) for BARR schools to use various strategies (e.g., posting their BARR experiences on social media, presenting at conferences, hosting other schools) to promote their experience with the model. More specifically, BARR Center believes that if more than half of schools that have

[34] We only conducted scale-up measurement for Years 2–4 for this strategy because BARR Center was not expected to make progress on these indicators in Year 1.

[35] Although BARR Center held a session on communication during their annual conference, attendance was not taken. However, the agenda confirms that this session took place.

implemented the model for one year were to use BARR communication strategies, they would significantly expand awareness of the model to other schools or districts.

For this strategy, about 13–20 percent of BARR-implementing schools used BARR communication strategies each year. A greater percentage of schools publicly communicated their experiences with implementing BARR in Year 3 (2019–20) and Year 4 (2020–21) than in Year 2 (2018–19). Given limitations in the available communications records, the percentage of BARR-implementing schools using BARR communication strategies may be greater than our calculations. We primarily used social media records and online news articles for our analysis, but schools use BARR communication strategies and expand awareness of the model in other ways. For example, BARR Center shared that schools have used the BARR website and monthly newsletter to learn more and connect one-on-one with BARR-implementing schools. In addition, schools engage in BARR Center's professional learning and thought leadership opportunities and share information with their networks.

In interviews with regional managers and coaches, participants reported that schools most frequently communicated about the BARR model through school site visits, social media, and school board presentations. Other means of expanding awareness of the BARR model included local news articles, district newsletters, and parent engagement events or communications. Furthermore, regional managers and coaches reported that school site visits and networking with local schools and districts were the most successful strategies used by their schools to increase awareness of the BARR model. When asked about additional resources that would help schools communicate about BARR, interviewees often mentioned the need for additional marketing materials, although some said that BARR Center has recently developed communication guidance for schools. Other areas of needed support included help with social media, standard processes for hosting site visits for schools interested in BARR, and opportunities to send more school and district staff to the BARR annual conference.

Strategies for Sustaining the BARR Model

Strategy: Sustain Cost of Services

As part of the i3 scale-up grant, participating study schools received grant funding for three years of BARR implementation. In Year 3 (2019–20) and Year 4 (2020–21) of the study, schools were eligible to seek out additional funding for their fourth and fifth years of BARR implementation. To scale the model beyond three years of implementation, BARR Center identified two strategies that would need to be in place. First, BARR Center would have to provide specific guidance about the cost of services to every school in their third year of implementation. Second, for the model to be sustainable, at least a third of schools in need of supplemental funding beyond the third year of BARR implementation would need to secure

philanthropic or third-party funding to help sustain BARR in subsequent years. The threshold was limited to one third of schools because BARR Center indicated that schools find it challenging to consider setting aside resources for a model they have been receiving free of charge for three years. BARR Center also noted that helping schools understand the costs associated with maintenance of the model is another challenge. BARR Center is building an accreditation process to encourage more schools to continue with programming; however, the accreditation process will not be fully in place until after the end of the grant award period (school year 2021–22).

BARR Center continues to make progress in supporting schools' efforts to sustain BARR implementation after grant funding ends. BARR Center shared costs of services/materials with nearly every school that had reached its third year of implementation and was still implementing the model. In Year 3 (2019–20), 71.4 percent of Cohort 1 schools that were interested in continuing BARR were able to successfully secure funding, which exceeded BARR Center's scale-up expectations. In Year 4 (2020–21), 42.1 percent of Cohort 1 and 2 schools were able to secure funding. This was a decrease from Year 3, which was impacted by COVID-related issues, but still exceeded the goals set by BARR Center.

Regional managers and coaches provided additional insights about the scale-up strategy of sustainable costs of services. During interviews we conducted with regional managers and coaches, they cited the following challenges regarding this strategy: (1) securing funds in smaller districts, (2) costs associated with the BARR coordinator position and block meetings, (3) challenging district approval processes, and (4) a lack of understanding about how BARR can be funded through various grants. Schools that were successful in securing funds often applied for state, federal, or philanthropy funding. However, regional managers and coaches still noted that it would be helpful for schools to see examples of how BARR can be funded and receive training in grant writing. BARR Center has worked to address this need by showing schools that BARR can fit with different funding opportunities, such as Title I funds or social-emotional program grants. BARR Center also created and shared a document and a "return on investment" calculator that demonstrated how BARR schools save money as a result of implementing BARR. This information was inserted into recruiting presentations and webinar slide decks. In some cases, BARR Center assisted schools in their search for funds by helping them write grant language and or by assisting them with budgeting tasks. Finally, when BARR Center became aware of funds that could be used by schools to purchase BARR, the BARR team provided that information directly to the schools.

Strategy: Sustain Training of Coaches at Scale

As the BARR model has expanded to more schools, BARR Center has increased the size of their team of coaches and trainers. These staff are responsible for sharing the BARR theory and practice with schools through individualized training and support. With such direct interactions

with schools, it is critical that coaches and trainers receive proper training from BARR Center. As part of its scale-up efforts, BARR Center aimed to have at least 84 percent of its active coaches and trainers be fully trained in Years 1–4 of the study. BARR Center exceeded this goal and, each year, had a team in which 90–95 percent of coaches and trainers had received training.

During the grant period, BARR Center modified its training for coaches to follow a mentoring model. In Years 1 and 2, new coaches were typically recruited from schools that had previously implemented the BARR model and were already familiar with BARR theory and practice. Upon hire, their onboarding consisted of direct mentoring from BARR Center's technical assistance director. In Years 3 and 4, however, BARR Center met the growing need for more coaches by expanding their recruitment and hiring coaches who were not affiliated with a school that had previously implemented the BARR model. Because these coaches did not have prior experience with the model, BARR Center created a formal mentoring structure that paired former and current BARR coaches (mentors) with new coaches (mentees). In this mentoring model, mentors joined their mentees' weekly calls with schools and their first three site visits to schools (at a minimum). Mentors provided feedback to new coaches and supported them in completing call notes, coaching fidelity forms, and coaching reports. This mentoring and supervised coaching continued until BARR Center's director of coaching and secondary schools accompanied the new coach on a site visit to ensure they were ready to coach without their mentor. In the summer of 2020, BARR Center further formalized the mentoring model by implementing a mentoring handbook, mentor training, and leveling system to differentiate support for new coaches based on their prior experiences.

BARR Center continues to develop mentoring and training opportunities so that coaches have adequate knowledge of the BARR model and align their coaching to its theory and practice. In anticipation of the model scaling to more schools and needing to hire more coaches, BARR Center developed a rubric that assesses coaches' proficiency in essential knowledge areas and coaching practices. With this rubric, BARR Center will be able to formally certify that coaches are providing high-quality support to schools and identify opportunities for additional training. During the grant period, BARR Center made progress on the development of the rubric by documenting coaching expectations, verifying the coaching rubric proof of concept, and piloting the rubric with longtime BARR coaches. BARR Center started using the coaching rubric during the 2021–22 school year.

In their interviews, regional managers and coaches reported that the most challenging aspects of their role included tailoring their support to schools and working with schools that were unresponsive to their communications. When asked how the training provided to coaches could be improved, interviewees said that they would like more opportunities to collaborate and standardize their coaching practices. In Year 3 (2019–20), regional managers and coaches

described efforts by BARR Center to provide more professional development and collaboration opportunities, such as mentoring, retreats, and book studies. Interviewees mentioned that they would appreciate more of these opportunities in the future.

Strategy: Sustain School-Level Staff Training at Scale

At the onset of BARR implementation, schools receive a two-day, on-site training for all educators involved in implementing the BARR model. The training provides an overview of the BARR model, gives staff an opportunity to practice implementing BARR strategies, and helps staff build community. Because this initial training provides foundational knowledge of the BARR model, BARR Center expected that successful scale-up of the model would require that 85% of BARR-implementing schools complete the two-day training in their first year of implementation. BARR Center exceeded this expectation in all four years of the study. Each year, 93–97 percent of BARR-implementing schools had received training. Across all cohorts, only three schools did not receive BARR training and it was due to the schools deciding not to proceed with BARR implementation.

BARR Center also offered additional trainings and professional development that were voluntary for schools to attend. In the second and third years of BARR implementation, schools are offered additional two-day trainings. Trainings cover different topics, such as trauma, substance abuse, effective team meetings, and equity. BARR Center also convened monthly PLC webinars for BARR coordinators and quarterly PLC webinars for school administrators. These events provided schools with an opportunity to discuss successes and challenges of implementation with BARR coaches and other schools implementing the BARR model.

Regional managers and coaches noted that school-level staff trainings generally went well and helped create buy-in for the BARR model. However, some schools faced challenges related to scheduling the training date and attendance. In particular, administrator attendance was a common challenge that emerged from the interviews. Although administrator attendance is not required by the BARR model, it is strongly encouraged. Regional managers and coaches also noted that schools faced challenges with paying teachers to attend the training during the summer, which is outside of the typical teacher contract. BARR Center has addressed these challenges by offering trainings later in summer and make-up trainings during the school year. Interviewees also reported that they provided information to absent teachers and staff by scheduling short meetings or leaving training materials for them. In Year 2 (2018–19) and Year 3 (2019–20), regional managers and coaches also reported that BARR Center revised their trainings to update the content and make virtual adaptations to address barriers caused by COVID-19.

In addition to participating in trainings, schools are asked to enter end-of-year data on student behavior and academic progress into a BARR database. In addition to reviewing real-time student data in block meetings and risk review, teachers and staff are asked to further build their data-driven practices by providing data to the BARR database to monitor their school's progress on achieving positive student outcomes and identify areas for improvement. Throughout the study, BARR Center faced challenges with schools providing data and did not receive data from 85 percent or more of schools each year, the expected threshold for this scale-up indicator. Instead, 40.9 percent of Cohort 1 treatment and control schools entered data in Year 2 (2018–19), 56.8 percent of Cohort 1 and 2 treatment and control schools entered data in Year 3 (2019–20), and 39.4 percent of Cohort 1, 2, and 3 treatment and control schools entered data in Year 4 (2020–21).[36] BARR Center reported that some barriers related to schools submitting data included timing of the data request, perceptions that the data request was optional, and lack of awareness about its usefulness, and possible competition with additional data collection efforts from AIR on student-level versions of the requested measures for this evaluation. After learning of these barriers, BARR Center took steps to adjust their outreach by sending the data request earlier and discussing the usefulness of annual reports with schools.

Strategy: Sustain School-Level Implementation

BARR Center expected schools to sustain the BARR model by implementing the eight BARR components with fidelity in each year of implementation. BARR coaches conducted annual site visits to each school and used fidelity rubrics to rate each school's implementation of the eight BARR components. The fidelity rubrics produced a "component score" for each BARR component that measured the extent to which schools implemented the BARR model. For each school, component scores were compared against predetermined component score thresholds to determine whether the school met model implementation expectations. To demonstrate sustainable implementation, BARR Center expected that 67 percent of schools participating in the model would meet or exceed component thresholds on four out of five priority components (restructuring, I-Time, block meetings, risk review, and professional development) and two out of three additional components (parent involvement, whole student emphasis, and contextual support) annually. BARR Center continues to make progress on this goal.

Thirty-three percent of BARR schools met the school-level fidelity threshold in Year 1 (2017–18), 55 percent of schools met the school-level fidelity threshold in Year 2 (2018–19), 35% of schools met the school-level fidelity threshold in Year 3 (2019–20), and 36 percent of schools

[36]No schools were expected to enter data in Year 1 (2017–18 school year).

met the school-level fidelity threshold in Year 4 (2020–21).[37] Schools were often more successful at implementing two out of the three additional BARR components (parent involvement, whole student emphasis, and contextual support) than implementing four out of the five priority components (restructuring, I-Time, block meetings, risk review, and professional development). In addition, schools showed mixed experiences with implementation over time. For example, some schools met adequate implementation for more components in their first year of implementation, and others saw an increase in the number of components meeting adequate implementation in later years of implementation.

Measuring Sustainable School-Level Implementation for Scale-Up Versus Measuring Fidelity of Implementation for Study Schools

As part of scale-up measurement, we examined all study schools' implementation of the BARR model over multiple years. We calculated school-level implementation using fidelity of implementation scores collected by BARR coaches during their annual site visits. This differs from the fidelity of implementation analysis presented in Chapter 2, which only focuses on treatment schools' first year of implementation and was calculated with implementation scores collected primarily by AIR researchers.

The regional managers and coaches we interviewed provided additional context about sustainable implementation of the BARR model. First, interview participants described several factors that they believe had an impact on schools' ability to implement BARR. These included a supportive BARR coordinator, supportive school administrators, ability to restructure the ninth grade, and contextual factors at the school (e.g., teacher turnover, number of high-need students). Regional managers and coaches reported that they adjusted their supports to schools by focusing on schools' individual concerns and tailoring their coaching to match schools' needs. Coaches also made themselves available whenever a school needed assistance. In Year 3, interview participants reported that they convened "task forces" to make virtual adaptations for block meetings, I-Time, and trainings to support schools during COVID-19-related school closures; revised BARR tracking spreadsheets and coaching reports; and created a resource guide for the back-to-school transition for school staff and administrators.

Each year, schools were most often able to meet implementation expectations for the professional development and contextual support components. Block meetings, whole student emphasis, and parent involvement were the most challenging components to implement. However, these findings do not completely align with the perceptions of regional managers and

[37] The scores we used to calculate school-level fidelity in Year 3 were collected by BARR coaches during their final site visit to schools during the 2019–20 school year before COVID-19 school closures. These site visits occurred between December 2019 and March 2020, which was midway through schools' first year of implementing the BARR model. In Year 4, BARR coaches were unable to measure implementation of one or more BARR components at 12 schools due to COVID-19. In these cases, implementation was rated as a "1" (i.e., "not yet implementing"), which may likely be a lower score than actual implementation.

coaches. When asked about the most successfully implemented BARR components prior to COVID-19 school closures, regional managers and coaches identified I-Time, block meetings, and whole student emphasis. Interviewees said that risk review and parent engagement were the most challenging components for schools prior to COVID-19. After school campus closures, regional managers and coaches said that block meetings and whole student emphasis were the most successfully implemented components, and I-Time was the most challenging component for schools.

During the grant period, BARR Center also planned to develop a rubric to formally certify schools that achieved adequate implementation of the BARR model. BARR Center achieved this goal by developing a school accreditation process in Year 3 and examining its proof of concept in Year 4. To be an accredited BARR school, schools must be in at least the third year of BARR implementation, ensure that 60% of BARR teachers have participated in BARR training, and meet with BARR Center staff to discuss accreditation. In addition, schools must achieve a designation of 2 or 3 (with 3 being the highest ranking) for implementation of three BARR strategies (I-Time, block meetings, and risk review), and schools must be able to identify improvement in student academics or social-emotional growth because of implementing the BARR model. BARR Center anticipates piloting the school accreditation rubric during the 2021–22 school year.

What were facilitators and barriers to successful scale-up of the BARR model?

The scale-up measurement analysis and interviews with regional managers and coaches provided insight into facilitators and barriers to scaling up the model. Overall, regional managers and coaches explained that the work of BARR Center staff is essential to the successful implementation and expansion of the BARR model. Regional managers and coaches support sustainable implementation of the model by coaching BARR coordinators and school staff. The high level of attention provided to individual schools continued during the school campus closures due to COVID-19, and schools were able to continue implementing the BARR model. One BARR coach shared the following about being seen as a valuable partner to schools during the 2019–20 school campus closures:

> *Specifically for coaching, I didn't change that much about what I did. I was still always present and available in supporting my school through the process, but I had more explicit conversations with schools about what I can do to support them and how I can help them and that it's okay if you need to vent about things, and it's okay if your questions aren't exactly about BARR. And so that was kind of nice because I think it was a turning point for some coordinators in really seeing us as support.*

Furthermore, approximately half of interviewed regional managers and coaches said that these supportive relationships help facilitate the recruitment of new BARR schools. Often, staff at current BARR schools help spread the word about the model or bring it to other schools. One coach explained:

> *…Call me any time. Ask any question. I will do whatever I can…I'm thinking back to the schools that are picking up [BARR] too. I think the fact that I'm really doing my best and I think I do a fairly good job at building strong relationships with the people I'm coaching that then they take it either with them to their new school or they are helping me spread it to other schools.*

In addition, regional managers and coaches frequently indicated that supportive school and district administration helps sustain BARR implementation. One BARR coach said:

> *For most, I would say the leadership at the school and the leadership at the district office [needs to be] on board. That is hands down the number one thing. At both of the schools that I coach, I have seen leadership change and seen buy-in go away because of that leadership changing, and it makes it really difficult to continue implementing with fidelity. So I would say by far, it's not just the site and it has to be the district as well.*

Another common facilitator to sustainable implementation of the model is the availability of flexible coaching supports. Regional managers and coaches frequently reported that they had to tailor their supports to schools' individual needs for implementation of the BARR model to move forward. For example, some schools experienced challenges with scheduling their school-level training date or having key school staff attend the training. To address these challenges, coaches and trainers worked with schools to schedule training dates later in the summer and found ways to provide information or training to those who were unable to attend. Once implementation began, coaches also had to adjust their supports to fit schools' individual contexts and provide ideas for how to best implement the model. One coach explained:

> *I talk a lot to schools about, "Okay. Here's the ideal of what I would like to see." Like, 'If you're doing the model to full fidelity, this is what it would look like.' However, schools are realistic places. They're not little models of what we would like to create. So, "Let me tell you what I would love to see if you were doing it, this, [with] whole fidelity, and then you tell me what the challenges are with that, and let's work together to see what we can create that's the closest to fidelity."*

As described in our discussion of the specific scaling strategies above, regional managers and coaches also identified barriers to successful scale-up of the BARR model. Related to expanding school and district awareness, regional managers and coaches reported that schools could use

more marketing materials and help with social media to communicate their experiences implementing the BARR model. This is aligned with our scale-up measurement analysis finding that schools did not communicate about BARR as often as BARR Center would have liked. However, interviewees noted that BARR Center has recently made efforts to provide more guidance in this area.

Furthermore, interviewees noted areas where existing BARR Center resources and processes could be streamlined to help with implementation and recruitment. Regional managers and coaches said that it would be helpful to create an organized database of BARR resources and templates that schools could access at any time. Interview participants also noted that it would be helpful to have a more structured process for coordinating site visits to "host schools" for schools that are interested in observing BARR activities and learning more about the model. BARR Center has begun to address these needs by creating a formal learning management system for BARR coordinators to access resources. Coaches and coordinators also can communicate with each other in this virtual space.

Summary

As described above, BARR Center identified six strategies they planned to use to grow and sustain the BARR model over time. The two strategies for growing the program to more schools included fortifying BARR Center infrastructure and expanding school and district awareness of the model. The four strategies used to ensure high-quality implementation as the model grew included sustaining the cost of service, coach training at scale, school-level staff training at scale, and school-level BARR implementation over time. We measured the extent to which BARR scale-up strategies were completed as planned by using BARR model data and measuring them against predetermined thresholds for each indicator of scale-up that were based on BARR Center's theoretical expectations and practical experiences.

BARR Center demonstrated progress on all six scale-up strategies during the grant period. BARR Center was particularly successful in scaling strategies that were primarily implemented by them. That is, BARR Center often met or exceeded expectations for fortifying infrastructure, establishing and sharing the cost of services, and scaling both school-level and coach trainings and supports. However, BARR scale-up strategies related to expanding school-level awareness and school-level implementation of the BARR model were less successfully scaled. These two strategies were primarily implemented by schools, not BARR Center. As such, BARR Center could only provide limited support to growing and sustaining those strategies. BARR Center recognizes this challenge and is working toward creating infrastructure (e.g., a school accreditation process) and supports (e.g., standardized coaching) to help schools increase the quality of their BARR model implementation and their efforts to expand awareness of the model.

Chapter 5: Study Limitations and Conclusion

Limitations

Study Design and Data Collection

The validity of the study findings for drawing conclusions about the impact of the BARR model depends on the integrity of the underlying experimental design as well as the proper representation of how consistently the model was implemented in the treatment schools. The randomization of 66 schools was successful in generating sets of treatment and control schools that were similar in school and student background characteristics. Only one Cohort 2 school dropped out of the evaluation, so the overall study design and its implementation do not present serious limitations for the interpretation of the findings.

However, four treatment schools decided not to implement BARR in their study year, and eight schools offered the BARR model only to some of their ninth-grade students even though they agreed to offer it to all their students when they first decided to participate in the study. This means that the evaluation's experimental impact estimates were likely attenuated relative to what they would have been if all schools and students in the treatment group had been exposed to BARR. This limitation especially affects Cohort 1 impact estimates. Note that the Complier Average Causal Estimates (CACE) introduced in Chapter 1 and presented in Chapter 4 attempt to address this limitation but can only do so under certain strict assumptions that may not be valid for all affected students in this study.[38]

Another limitation associated with the overall study design is that estimates of BARR's impact on teachers and school climate depended on teacher self-report on a survey administered at the end of each school year (or in the following fall for the 2019–20 cohort). Such self-reports do not always reliably capture the underlying constructs. Moreover, there is some degree of survey nonresponse in these teacher surveys, which may affect the impact estimates.

Survey nonresponse was a significant problem for the student surveys we administered to capture impacts on student experiences and school climate (see Appendix F). As a result, we decided not to include impact estimates from these student surveys in this report. This limits our assessment of BARR's impact on the school environment to the perspective of the teachers presented in Chapter 4 and the qualitative implementation data presented in Chapter 3.

[38] The most important of these assumptions is that students in schools assigned to BARR who did not receive BARR services are assumed not to be impacted by their assignment to BARR. This assumption may be violated, for example, if structural changes were started in a treatment school and then abandoned when school leadership decided to postpone BARR implementation by a year. The resulting disruption could have negative consequences for students in that school without the expected attendant benefits of exposure to BARR.

There was also some student attrition in the PSAT achievement test outcomes. This factor also introduces a risk of nonresponse bias for these outcomes. Because the COVID-19 pandemic made test administration for Cohort 3 very difficult, PSAT testing rates were particularly low for this cohort. For this reason, we decided not to conduct PSAT impact analyses for Cohort 3. Further, for Cohorts 1 and 2, we only included schools that administered the PSAT/NMSQT in the fall of 10th grade in the impact analyses. As a result, our cross-cohort impact estimates on the PSAT outcomes are statistically less powerful than intended in the original study design.

Finally, a limitation related to the study design is that the impact and implementation evaluation focus only on the first year of implementation of BARR in treatment schools. As noted in Chapter 1, BARR is a three-year model. That is, by the time we measured its impacts, it was not yet fully implemented. The reason for this design constraint is that, during the recruitment of schools for the study, it was considered impossible to exclude control schools from BARR for more than one year. As a result, our ability to contrast outcomes in BARR schools with outcomes in control schools ended after Year 1. To the extent that BARR's effectiveness grows as it is more fully implemented, the impact estimates presented here likely underestimate the true effects of BARR. (This is also the case for prior evaluations of BARR; see e.g., Borman et al., 2021).

Delayed and Long-Term Impacts

Related to the fact that this evaluation only had a one-year treatment contrast, it is also important to note that effects of interventions that change the organization and operation of schools are often delayed (e.g., Eastwood & Louis, 1992; Gleason et al., 2019). Those delayed effects can occur at the school, teacher, and student levels and can make one-year impact estimates premature. Our study did not include long-term follow-up of students, teachers, and schools, so we do not know whether BARR impacted teacher and student outcomes much beyond ninth grade or beyond the initial implementation year. We also do not know from this study how BARR impacts long-term student outcomes such as high school graduation or college enrollment, even though these outcomes are part of the BARR model's logic model.

The COVID-19 Pandemic

The COVID-19 pandemic, a once-in-a-lifetime disruption, compromised BARR model implementation in Cohort 3 and many of the study activities for that cohort as well. The most important limitation this imposes on the evaluation is that Cohort 3 findings cannot be fully generalized beyond that specific context. We tried to minimize the impact of the pandemic on the study by using pre-pandemic data for Cohort 3 wherever possible (i.e., data from the first half of the 2019–20 school year), but that approach introduced new limitations as student experiences and outcomes in the first half of a school year do not necessarily reflect what happens in the second half. Also, some data collection activities (specifically, the student and

teacher surveys and school staff interviews) were delayed to the tenth-grade school year. Because of this, it is important to interpret all Cohort 3 findings with caution.

Conclusion

This is an evaluation of the BARR model as it was implemented at scale. This means that, compared to the two prior evaluations of the BARR model, BARR Center was developing processes to allow for less direct involvement of the developer and her team in the day-to-day operation of the BARR model. These operational changes may have contributed to the less robust fidelity of implementation we found in this study compared to previous studies, as reported in Chapter 2. In addition to the four schools that did not implement BARR at all during the study year, and the schools that only offered it to part of their ninth-grade class, there were other schools across all cohorts that did not implement several of the eight BARR model components with fidelity. Because of this, the evaluation results for an average school in our evaluation sample do not represent what a school that is fully committed to the BARR model and ready to implement it might experience.

BARR Center demonstrated considerable progress in scaling up its model to reach many more schools than was possible during the earlier development and validation stages of the BARR model. Specifically, BARR Center met or exceeded scale-up expectations for developing sustainable infrastructure, determining and sharing cost of services, and scaling school-level and coach trainings and supports. However, components related to school-level awareness and expansion and school-level implementation were more challenging for the BARR Center to scale. Implementation of these two components were primarily the responsibility of schools, not BARR Center. As such, BARR Center could only provide limited support to scaling those components.

The implementation data highlight that schools had the most success implementing professional development and risk review meetings and the least success implementing block meetings. Our interview data provided some additional context for understanding why this might be. For the professional development component, even though some schools reported challenges associated with scheduling, the foundational training and BARR coaching visits were reported as the most beneficial professional development BARR schools received. Schools may have prioritized the implementation of this component as they found it most valuable during the first year of BARR implementation. By contrast, schools struggled most with implementing block meetings consistently. Respondents reported that insufficient time and the frequency of meetings were challenges to implementing these meetings, which may partly explain the difficulty of implementing them with fidelity.

Schools in all three cohorts experienced implementation challenges. Overall, only 50 percent of assigned treatment schools met adequate school-level implementation, suggesting that implementing the BARR model is challenging for many schools, at least in the first year of implementation and that sustained support and commitment are needed for successful adoption of the model.

Our impact analyses show that assignment to BARR improved teachers' collaboration with their peers, their data use, and a range of other teacher outcomes. These positive effects on teacher-reported outcomes occurred in all three cohorts and persisted even after schools were closed for much of the 2019–20 school year due to the COVID-19 pandemic.

Impact findings on student achievement and course performance are consistent with those we found in our 2019 i3 validation study of BARR (Borman et al., 2021; Bos et al., 2019), and Corsello and Sharma found in their 2015 evaluation of BARR's first i3 development grant. All three studies found substantial positive impacts on credit attainment and grades, as well as reductions in course failure, and smaller effects on academic achievement. In this evaluation, we also found favorable effects on absenteeism and suspensions. All these effects were more substantial when we adjusted our impact estimates for the fact that some schools and many students in the BARR treatment group did not receive BARR in the study year.

There was considerable variation in impacts across different groups of students, with traditionally lower performing groups typically benefiting more from having the BARR model in their school. Impacts also varied across the three cohorts of schools. This variation may reflect underlying differences in the composition of these three cohorts, differences in model implementation, and/or differences due to the COVID-19 pandemic, which significantly impacted Cohort 3.

The findings from this evaluation therefore support the conclusion we made at the end of the validation study, which is that BARR is an effective model for schools aiming to improve students' transitions into ninth grade, reduce course failure, and narrow existing gaps in student academic outcomes between different demographic groups of students. This conclusion is even more warranted considering the dampening effect of some BARR treatment schools not delivering the intervention, as well as the COVID-19 pandemic.

References

Abe, Y., & Gee, K. A. (2014). Sensitivity analyses for clustered data: An illustration from a large-scale clustered randomized controlled trial in education. *Evaluation and Program Planning*, *47*, 26–34.

Allensworth, E. M. (2013). The use of ninth-grade early warning indicators to improve Chicago schools. *Journal of Education for Students Placed at Risk*, *18*(1), 68–83.

Allensworth, E. M., & Easton, J. Q. (2005). *The on-track indicator as a predictor of high school graduation.* University of Chicago Consortium on Chicago School Research. https://consortium.uchicago.edu/sites/default/files/2018-10/p78.pdf

Alvidrez, J., & Weinstein, R. S. (1993). The nature of "schooling" in school transitions: A critical re-examination. *Prevention in Human Services, 10*, 7–26.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4),

561–573.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using

instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455.

Beltman, S., Mansfield, C., & Price, A. (2011). Thriving not just surviving: A review of research on teacher resilience. *Educational Research Review*, *6*(3), 185–207.

Bempechat, J., & Shernoff, D. J. (2012). *Parental influences on achievement motivation and student engagement.* In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 315–342). Springer.

Benner, A., Boyle, A., & Bakhtiari, F. (2017). Understanding students' transition to high school: Demographic variation and the role of supportive relationships. *Journal of Youth and Adolescence*, *46*, 1–14.

Benner, A. D., & Graham, S. (2009). The transition to high school as a developmental process among multiethnic urban youth. *Child Development, 80*(2), 356–376.

Borman, T. H., Bos, J. M., Park, S. J., & Auchstetter, A. (2021). Impacting 9th grade educational outcomes: Results from a multisite randomized controlled trial of the BARR model. *Journal of Research on Educational Effectiveness, 14*(4), 812–834.

Bos, J. M., Dhillon, S., & Borman, T. (2019). *Building Assets and Reducing Risks (BARR) validation study: Final report.* American Institutes for Research. https://eric.ed.gov/?id=ED602462

Connell. J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self processes in development: Minnesota Symposium on Child Psychology, 29*, 244–254. Lawrence Erlbaum Associates Inc.

Cook, C., Fowler, H., & Harris, T. (2008). *9th grade academies: Easing the transition to high school.* North Carolina Department of Public Instruction. https://hjackson.weebly.com/uploads/2/2/4/3/22436168/dpi_9thgradeacademies.pdf

Corrin, W., Sepanik, S., Rosen, R., & Shane, A. (2016). *Addressing early warning indicators: Interim impact findings from the Investing in Innovation (i3) evaluation of Diplomas Now*. MDRC. https://files.eric.ed.gov/fulltext/ED566904.pdf

Corsello, M., & Sharma, A. (2015). *The Building Assets-Reducing Risks program: Replication and expansion of an effective strategy to turn around low-achieving schools* (i3 Development Grant, Final Report). https://files.eric.ed.gov/fulltext/ED560804.pdf

Eastwood, K. W., & Louis, K. S. (1992). Restructuring that lasts: Managing the performance dip. *Journal of School Leadership*, *2*(2), 212–224.

Eccles, J. S., & Roeser, R. W. (2011). *School and community influences on human development*. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental science: An advanced textbook* (pp. 571–643). Psychology Press.

Ellerbrock, C. R., & Kiefer, S. M. (2013). The interplay between adolescent needs and secondary school structures: Fostering developmentally responsive middle and high school environments across the transition. *The High School Journal, 96*(3), 170–194.

Faria, A.-M., Sorensen, N., Heppen, J., Bowdon, J., Taylor, S., Eisner, R., & Foster, S. (2017). *Getting students on track for graduation: Impacts of the Early Warning Intervention and Monitoring System after one year* (REL 2017–272). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. https://files.eric.ed.gov/fulltext/ED573814.pdf

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, *95*(1), 148–162.

Gleason, P., Crissey, S., Chojnacki, G., Zukiewicz, M., Silva, T., Costelloe, S., & O'Reilly, F.

(2019). *Evaluation of support for using student data to inform teachers' instruction* (NCEE 2019-4009). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/pubs/20194008/pdf/20194008.pdf

Halbesleben, J. R. (2006). Sources of social support and burnout: A meta-analytic test of the conservation of resources model. *Journal of Applied Psychology*, *91*(5), 1134–1145.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis.

Psychological Methods, 3(4), 486–504.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1,* 221–233. University of California Press.

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, *79*(1), 491–525.

Johnson, S. M., Kraft, M. A., & Papay, J. P. (2012). How context matters in high-need schools: The effects of teachers' working conditions on their professional satisfaction and their students' achievement. *Teachers College Record*, *114*(10), 1–39.

Juvonen, J., Espinoza, G., & Knifsend, C. (2012). The role of peer relationships in student academic and extracurricular engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 387–401). Springer.

Kawulich, B. B. (2004). Data analysis techniques in qualitative research. *Journal of Research in Education*, *14*(1), 96–113.

Kontos, S., & Wilcox-Herzog, A. (1997). Teachers' interactions with children: Why are they so important? Research in review. *Young Children, 52*(2), 4–12. https://eric.ed.gov/?id=EJ538100

Liem, G. A., & Martin, A. J. (2011). Peer relationships and adolescents' academic and non-academic outcomes: Same-sex and opposite-sex peer effects and the mediating role of school engagement. *British Journal of Educational Psychology*, *81*, 183–206.

Maas, C. J., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, *58*(2), 127–137.

Mahan, P. L., Mahan, M. P., Park, N. J., Shelton, C., Brown, K. C., & Weaver, M. T. (2010). Work environment stressors, social support, anxiety, and depression among secondary school teachers. *AAOHN Journal*, *58*(5), 197–205.

Mansour, M., & Martin, A. J. (2009). Home, parents, and achievement motivation: A study of key home and parental factors that predict student motivation and engagement. *Australian Educational and Developmental Psychologist*, *26*, 111–126.

Martin, A. J., & Collie, R. J. (2019). Teacher–student relationships and students' engagement in high school: Does the number of negative and positive relationships with teachers matter?. *Journal of Educational Psychology*, *111*(5), 861.

Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research, 79*(1), 327–365.

Neild, R. C. (2009). Falling off track during the transition to high school: What we know and what can be done. *Future of Children, 19*(1), 53–76.

Neild, R. C., Stoner-Eby, S., & Furstenberg, F. (2008). Connecting entrance and departure: The transition to 9th grade and high school dropout. *Education and Urban Society, 40*(5), 543–569.

Osher, D., Cantor, P., Berg, J., Steyer, L., & Rose, T. (2018). Drivers of human development: How relationships and context shape learning and development. *Applied Developmental Science*, *22*(1), 1–31.

Pomerantz, E. M., & Moorman, E. A. (2010). Parents' involvement in children's schooling: A context for children's development. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development.* Routledge.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (Exp. ed.).

University of Chicago Press.

Raudenbush, S.W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). SAGE Publications.

Reyes, O., Gillock, K. L., Kobus, K., & Sanchez, B. (2000). A longitudinal examination of the transition into senior high school for adolescents from urban, low-income status, and predominantly minority backgrounds. *American Journal of Community Psychology*, *28*(4), 519–544.

Roderick, M., Kelley-Kemple, T., Johnson, D. W., & Beechum, N. O. (2014). *Preventable failure: Improvements in long-term outcomes when high schools focused on the 9th grade year*. University of Chicago Consortium on Chicago School Research.

Ronfeldt, M., Farmer, S. O., McQueen, K., & Grissom, J. A. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal*, *52*(3), 475–514.

Ruzek, E. A., Hafen, C. A., Allen, J. P., Gregory, A., Mikami, A. Y., & Pianta, R. C. (2016). How teacher emotional support motivates students: The mediating roles of perceived peer relatedness, autonomy support, and competence. *Learning and Instruction*, *42*, 95–103.

Scales, P. C., Pekel, K., Sethi, J., Chamberlain, R., & Van Boekel, M. (2020). Academic year changes in student-teacher developmental relationships and their linkage to middle and high school students' motivation: A mixed methods study. *The Journal of Early Adolescence*, *40*(4), 499–536.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton Mifflin.

Sibley, E., Theodorakakis, M., Walsh, M. E., Foley, C., Petrie, J., & Raczek, A. (2017). The impact of comprehensive student support on teachers: Knowledge of the whole child, classroom practice, and teacher support. *Teaching and Teacher Education*, *65*, 145–156.

Styron, R., & Peasant, E. (2010). Improving student achievement: Can 9th grade academies make a difference? *International Journal of Education Policy and Leadership*, *5*(3).

Teven, J. J., & McCroskey, J. C. (1997). The relationship of perceived teacher caring with student learning and teacher evaluation. *Communication Education*, *46*, 1–9.

U.S. Department of Education. (2019). *2019 Mathematics Assessment.* Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). https://nces.ed.gov/nationsreportcard/mathematics/

U.S. Department of Education. (2017). *2017 Mathematics Assessment.* Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). https://nces.ed.gov/nationsreportcard/mathematics/

U.S. Department of Education. (n.d.a.). Four-year adjusted-cohort graduation rates—School year 2016–17. ED*Facts* Data Documentation. http://www.ed.gov/edfacts

U.S. Department of Education. (n.d.b.). Four-year adjusted-cohort graduation rates—School year 2017–18. ED*Facts* Data Documentation. http://www.ed.gov/edfacts

U.S. Department of Education. (n.d.c.). Four-year adjusted-cohort graduation rates—School year 2018–19. ED*Facts* Data Documentation. http://www.ed.gov/edfacts

Uvaas, T., & McKevitt, B. C. (2013). Improving transitions to high school: A review of current research and practice. *Preventing School Failure*, *57*(2), 70–76.

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, *24*(1), 80–91.

What Works Clearinghouse. (2020). *What Works Clearinghouse Standards Handbook, Version 4.1.* U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://eric.ed.gov/?id=ED602036

White, H. (1980). *A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica, 48*(4), 817–830.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement.* MESA Press.

Zeileis, A., Köll, S., & Graham, N. (2020). Various versatile variances: An object-oriented

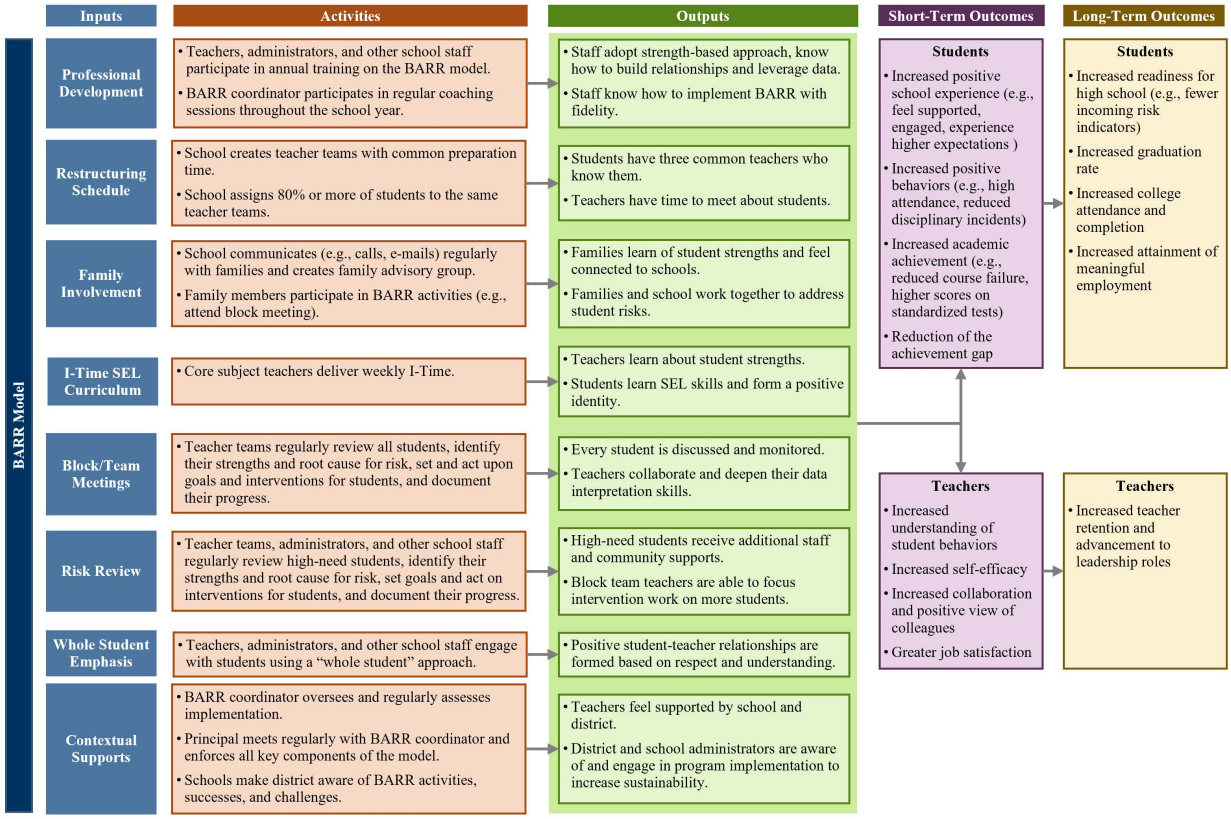
implementation of clustered covariances in R. *Journal of Statistical Software, 95*(1), 1–36.

Appendix A. Evaluation Design

A1. BARR Logic Model

This section provides a graphic representation of the BARR logic model.

Exhibit A1. BARR Logic Model



A2. School Recruitment, Selection, and Randomization

This school-level cluster randomized controlled trial includes three cohorts of schools studied in the 2017/18, 2018/19, and 2019/20 school years. The spring prior to each of these school years, BARR Center recruited schools to participate in the study. To participate, schools had to agree to random assignment and to facilitating data collection for the evaluation. The BARR Center focused on recruiting high schools with a traditional 9–12 grade structure where most ninth-grade students did not meet proficiency standards on state math and English language arts (ELA) assessments, and/or most students were from traditionally disadvantaged groups (e.g., students eligible for free or reduced-price lunch [FRPL]).

BARR Center's recruitment pool for Cohort 1 (2017/18) was limited to the lowest performing schools across the nation with students at high risk for poor educational outcomes. During the 2017/18 school year, BARR Center noted that many of the recruited Cohort 1 schools did Appendix A.not have the basic infrastructure needed to implement the model. As a result, BARR Center spent much of their coaching time with these schools helping treatment schools create structures to implement the model and support basic student needs.

BARR Center used this information to establish minimum infrastructure requirements that schools needed to meet to implement the model and be selected for Cohort 2 and Cohort 3. These additional requirements reduced the potential pool of schools that could be recruited. Consequently, although Cohorts 2 and 3 included low-performing schools with high-need students, the schools included in these two cohorts had the infrastructure needed to implement BARR during the study year for their cohort.

BARR Center recruited schools on a rolling basis during the winter and spring prior to each study year. To maximize the time available to the BARR Center for helping schools assigned to the BARR treatment group restructure their schedules, a key component of the model, the study team randomly assigned recruited schools to treatment or control status in blocks as they became available. That is, the AIR team did not wait for all the necessary schools to be recruited before beginning randomization. The randomization blocks were thus determined by the region where a school was located (i.e., Appalachia & Midwest, Southwest, Northeast) and the timing of recruitment. Random assignment ratios varied by block due to differences in the number of schools recruited within each region and BARR Center's attempt to manage the number of schools that would receive BARR services in each region. Exhibit A2 describes the random assignment blocks by cohort, including the random assignment ratio for each block.

Exhibit A2. Block Randomization Type Used Across Cohorts

Randomization Type (Ratio of Treatment [T] to Control [C])	Number of Blocks			Total Number of Blocks	Total Number of Schools
	Cohort 1	Cohort 2	Cohort 3		
Block 1T:1C	1	5	11	17	34
Single-school assignment*	2	3		5	5
Block 1T:2C		1		1	3
Block 2T:1C	3	2		5	15
Block 3T:1C	1			1	4
Block 3T:2C	1			1	5
Total	**8**	**11**	**11**	**30**	**66**

* One Cohort 1 single-school random assignment was weighted for 1T/2C. All other single-school random assignments were weighted for 1T/1C.

Treatment schools were assigned to implement the BARR model in the school year following recruitment (i.e., 2017/18, 2018/19, or 2019/20). Control schools were assigned to conduct business as usual during the study year and then invited to implement the BARR model in the year following the study year. In total, the study includes 66 schools (37 treatment and 29 control) across three cohorts, with all students entering ninth grade in the sample schools during the study year.

A3. Data Sources

We used a variety of data sources to answer the research questions regarding the scaling, implementation, and impact of BARR. This section details the data sources used in this study. Instruments referenced in this section can be found in Appendix B.

Some Cohort 3 data collection efforts were delayed or canceled due to the COVID-19 pandemic, which caused widespread school closures in spring 2020. All Cohort 3 schools participating in the study either transitioned to virtual learning in March 2020 or elected to end the school year early.[39] These closures caused most student testing to be canceled in spring 2020 and impacted attendance, discipline, and grade-level promotion policies.

To understand scaling, implementation, and impacts of the BARR model during this time, we added specific questions to Cohort 3 interview protocols and surveys. We also collected information on changes in school policies following Cohort 3 school closures to inform the inclusion of Cohort 3 administrative data in the analysis. In cases where data quality was compromised, we excluded certain administrative data sources or used proxy measures.

BARR Scale-up and Sustainability Data Sources

We conducted semi-structured interviews with BARR regional managers and BARR coaches at the end of each study year to understand the facilitators of and barriers to successful scale-up and sustainability of the BARR model during the grant period. The BARR coach interviews focused on BARR infrastructure, school and district awareness of the model, sustainable cost of services, coach training, school-level staff training, and school-level BARR implementation. The interview protocols for Cohort 3 schools also included questions about whether and how the COVID-19 pandemic affected coaching and school-level implementation. We conducted four 90-minute phone interviews in fall 2018, five interviews in summer 2019, and five interviews in summer 2020. Coaches were eligible to participate in interviews if they had provided coaching to one or more schools during the previous school year, and interviewees were randomly selected from the list of eligible coaches. Because coaches could oversee schools during each year of the study, there were some coaches who participated in interviews more than once. We

[39] One school in Cohort 3 closed two weeks early, and another indicated that their last day of classes was March 11, 2020.

recorded interviews with the permission of BARR regional managers and BARR coaches, and transcribed the recordings. If we were unable to record an interview, we took detailed notes on the interview responses for analysis.

To understand the extent to which BARR scale-up strategies were carried out as planned during the grant award period, we met annually with BARR Center staff to document their plans to grow and sustain the BARR model during and after the grant award period. These conversations were the basis for the scale-up measurement framework. We designed the framework by first asking BARR Center staff to identify critical barriers to growing and sustaining the BARR model. We then translated these barriers into strategies (conditions) needed to grow the BARR model over time and sustain long-term positive BARR model outcomes for schools participating in the BARR model. These strategies included BARR infrastructure, school and district awareness of the model, sustainable cost of services, coach training at scale, school-level staff training at scale, and sustainable school-level implementation. Appendix G provides additional details on the indicators measured for each of these six scale-up strategies.

We requested communication records, school and coach training attendance records, staffing information, and BARR-collected assessments of school BARR model implementation from BARR Center and Hazelden Publishing[40] (for between one and four years depending on the scale-up measure) to understand what steps BARR Center took to grow and sustain the model during the grant award period.

BARR Model Implementation Data Sources

We conducted semi-structured interviews with teachers, administrators, and (when applicable) BARR coordinators at each school to understand the facilitators of and barriers to model implementation. We recorded interviews with the permission of respondents and transcribed the recordings. If we were unable to record an interview, we took detailed interview notes for analysis.

The BARR teacher interview protocol focused on the eight BARR components, and the interview protocol for control teachers focused on the use of similar strategies. We also asked teachers whether there had been any major internal or external challenges that year, any challenges specific to BARR, and their perception of ninth-grade students' experiences that year. The interview protocols for Cohort 3 schools included additional questions about whether and how the COVID-19 pandemic affected their school.

We conducted site visits in spring 2018 with the Cohort 1 treatment schools; interviewed 50 teachers, 12 administrators, and 12 BARR coordinators; and observed meetings and classes in all the visited schools. In spring 2019, we conducted site visits with the Cohort 2 treatment

[40] BARR Center partners with Hazelden Publishing to hire and train BARR coaches. These coaches work directly with school BARR coordinators placed in schools implementing the BARR model.

schools and interviewed 51 teachers, 10 administrators, and 10 BARR coordinators (Exhibit A3). For Cohort 3, we canceled site visits to treatment schools due to COVID-19 restrictions. We conducted phone interviews with teachers, administrators, and BARR coordinators at 10 of the 11 Cohort 3 treatment schools in fall 2020 (the start of the next school year). Seven of the 11 Cohort 3 control schools also participated in phone interviews. One Cohort 3 treatment school and two Cohort 3 control schools were not able to participate in fall 2020 interviews due to delays in district research approvals related to COVID-19. Two additional Cohort 3 control schools declined to participate, citing limitations on availability related to COVID-19.

Exhibit A3. Interviews at Treatment Schools by Cohort[41]

	Number of Interviews at Treatment Schools		
Cohort	Teacher	Administrator	BARR Coordinator
1	50	12	12
2	51	10	10
3	27	9*	10

* One BARR coordinator also was an administrator for the school, but in this exhibit, we are counting this interview as a BARR coordinator interview.

To understand the extent to which the BARR model was implemented as intended in treatment schools during the study year, we observed the implementation of three components (block meetings, I-Time lessons, and risk review meetings) of the BARR model in treatment schools using fidelity rubrics that were first developed for the 2010 Investing in Innovation (i3) development project and subsequently revised for the 2014 i3 validation study. We also interviewed BARR coordinators at treatment schools who coordinated and supported BARR implementation during the study year. During structural component review interviews, the BARR coordinators shared and explained their ratings on several indicators associated with the implementation of the eight key components of the BARR model. These activities took place during annual spring site visits for Cohorts 1 and 2. Because of school closures due to COVID-19, we were unable to conduct site visit observations of BARR activities in Cohort 3. We measured implementation fidelity for Cohort 3 using ratings from interviews conducted with BARR coordinators in fall 2020 and the most recent implementation fidelity scores collected by BARR staff during the 2019/20 school year coaching site visits.

[41] We also conducted interviews at control schools, namely, 42 teacher interviews and 10 administrator interviews in Cohort 1, 45 teacher interviews and 11 administrator interviews in Cohort 2, and 17 teacher interviews and 7 administrator interviews in Cohort 3.

Ratings for each fidelity indicator from the BARR coordinator structural review interview and the observed BARR activities were assigned on a scale of 1 to 7, which correspond to the following levels of implementation:

- **Not Yet (1–2):** Activity has not been implemented or has been implemented to a limited extent (e.g., very preliminary planning or discussions have occurred, but no concrete planning is in place; activities have occurred sporadically, if at all, and to less than one third the expected frequency for full implementation). In the case of shifts in practice, this is occurring with few, if any, teachers or staff.
- **Emerging (3–5):** Activity has been implemented in an emerging manner so that concrete planning and, in some cases, activities have occurred, but not to the full extent as intended for the BARR model (e.g., planning has begun and concrete details, such as schedules or processes, have been established and finalized; some activities have been implemented, but they represent only a small proportion of the total number of activities that should be in place, such as monthly meetings rather than weekly meetings). In the case of shifts in practice, this may occur with some, but not all, teachers or staff, or it may occur with all teachers or staff, but at a superficial or beginning level.
- **In Place (6–7):** Activity has been implemented to a high degree, with all or nearly all anticipated activities completed as intended. In the case of shifts in practice, this has occurred with all or nearly all teachers or staff and has occurred almost completely to the extent intended.

Data Sources to Assess BARR Impact on Teachers and Students

Teacher Survey

To capture ninth-grade teachers' perception of self, students, and schools in BARR and control groups, we administered a survey to core subject teachers in both groups who taught ninth-grade students during the study year. Teachers of BARR and control group students received and completed the same survey instrument. The teacher survey was administered in spring 2018 and spring 2019, respectively, for Cohort 1 and Cohort 2. The administration of the Cohort 3 teacher survey was delayed to fall 2020 due to COVID-19-related school closures. The following constructs were measured:

- **Strength-based mindset**. Teachers are aware of students' strengths and interests and know how to use these strengths and interests to support and guide student engagement in the classroom.
- **Postsecondary educational expectations.** The extent to which teachers believe students will be prepared for and participate in postsecondary education opportunities.

- **Student academic motivation.** The extent to which teachers perceive their students to be motivated to learn.
- **Perception of students' behavior, commitment, and attitudes.** The extent to which teachers expect students to help their peers and make the right choices in a given situation.
- **View of students' observed behavior, commitment, and attitudes.** The extent to which teachers observe students acting in ways that support positive peer relationships, achieving academic objectives, and demonstrating the motivation to succeed.
- **Use of data.** Teachers have access to data, discuss data, and use data to differentiate instruction for students.
- **Teacher self-efficacy.** Teachers believe that they have the ability and power to affect learning, motivation, and behavior in classrooms.
- **Collaboration with and view of colleagues.** Teachers work together, trust each other, and have shared responsibilities and teaching approaches.
- **View of school supports provided to teachers.** Teachers feel that the school provides students and staff with the necessary resources and support services.

In addition, Cohort 3 teachers were asked questions about classroom educational expectations and access to technology during school closures as well as teacher efficacy and comfort with technology. The following two technology constructs were measured:

- **Comfort with technology.** The extent to which teachers are comfortable using a variety of classroom technologies to support student engagement and learning.
- **Self-efficacy for instructional use of technology.** Teachers believe that they have the ability and power to affect learning, motivation, and behavior using current educational technologies.

Student Survey

To measure students' experiences in school, we administered a survey to students in all study schools in the spring of ninth grade for Cohorts 1 and 2, and in the fall of tenth grade for Cohort 3. This survey was adapted from similar surveys in the two earlier impact evaluations of BARR (Bos et al., 2019; Corsello & Sharma, 2017). It measured student experiences and attitudes along the following six constructs: supportive relationships, expectations & rigor, engagement, emotional safety & school climate, future orientation, and self-efficacy.

Unfortunately, the administration of this survey resulted in lower response rates than expected. As detailed in Appendix F, response rates for the survey were quite low, especially in Cohort 3. Response rates also varied considerably by school and by research group and cohort. Moreover, we found evidence of nonresponse bias when we compared administrative

outcomes for respondents and nonrespondents. As described in Appendix F, we were unable to use imputation or other mitigation techniques for missing data and survey nonresponse due to being unable to match the student survey data to other student-level data in many of the study schools. (This was due to the masking of student identifiers in administrative data, which is increasingly common in education research.) Because of these data quality and bias concerns, we decided not to include student survey impact analyses in the body of the report. Additional details about the student survey data and this decision are in Appendix F.

Student Administrative Data

To examine ninth-grade students' academic outcomes (i.e., passing all core courses, PSAT scores, GPA) and their behaviors (i.e., chronic absenteeism, suspensions, persistence to 10th grade) in the BARR and control groups, we requested administrative data from study schools, districts, and, in some cases, state education agencies. Completeness and quality of these administrative data was a notable concern for Cohort 3. School closures in spring 2020 resulted in many changes in attendance, testing, grades, and credit award policies and procedures. As mentioned earlier, to salvage as much Cohort 3 data as possible, we collected information on changes in school policies following Cohort 3 school closures. In cases where the quality of these data was too compromised, we focused on pre-COVID-19 measures (i.e., for the first half of the 2019/20 school year) or used proxy measures. We measured the following outcomes:

- **Core course credits earned in ninth grade.**[42] Grade 9 students' average course credit attainment for all core courses (i.e., ELA, math, science, social studies) attempted during the school year
- **Passing all core courses in ninth grade.** Proportion of grade 9 students who passed all core courses (i.e., did not fail one or more core courses) they were enrolled in during the school year
- **GPA at the end of ninth grade.** Grade 9 students' average GPA at the end of the school year
- **Performance on ELA PSAT in the fall of 10th grade.**[43] Grade 10 students' average PSAT (or equivalent test[44]) evidence-based reading and writing section score

[42] In Cohort 3, only schools where it was possible for students to lose credit (i.e., students were not automatically given credit) after COVID-19-related school closures were included in the analysis.

[43] School closures limited our PSAT test administration in fall 2020. Sixteen schools were unable to administer the test due to various reasons associated with COVID-19, such as being in a remote learning mode or having insufficient staff to maintain social distancing requirements during testing. In the six schools that administered the PSAT or a similar test (i.e., one school administered the PreACT) during the 2020/21 school year, test administration was not systematic (e.g., offered on a day most students had remote learning) and likely biased. Consequently, Cohort 3 schools were excluded from the PSAT analyses.

[44] Of the 2017/18 cohort of schools, three schools administered the PSAT 10 in spring 2018. Of the 2018/19 cohort of schools, one school administered the PSAT 10 in spring 2019 ,and one school administered a released copy of the PSAT under standard proctoring procedures in fall 2019.

- **Performance on Math PSAT in the fall of 10th grade.** Grade 10 students' average PSAT (or equivalent test) evidence-based math section score
- **Chronic absenteeism in ninth grade.**[45] Proportion of grade 9 students who missed 10 percent or more of enrolled school days during the school year
- **Suspensions in ninth grade.** Proportion of grade 9 students who received one or more in-school or out-of-school suspension during the school year
- **Persistence to 10th grade.** Proportion of grade 9 students who enrolled in 10th grade in the same school the following year

A4. Measurement

As noted in the previous section, we used a variety of data sources to answer the research questions regarding the implementation, impact, and scale-up of BARR. This section details the measurement approaches used in this study.

Measurement of Facilitators and Barriers to Implementation

We used the qualitative analysis software NVivo to code the transcripts and interview notes from BARR schools for each year of data collection.[46] We established interrater reliability across the teams by having analysts code the same interview transcript and using NVivo to calculate a Kappa coefficient for each analyst.[47] All analysts produced a Kappa coefficient greater than 0.75 with the lead qualitative analyst, which indicates excellent agreement.

We used multiple analytic methods to systematize our review of the data, including qualitative content analysis and constant comparison (Kawulich, 2004). We coded these interview data using a coding schema containing *a priori* codes based on the eight components of BARR. A subset of analysts then reviewed the coded data from BARR schools to identify the facilitators of and barriers to implementing each component of BARR. In this report, we present the most frequently discussed facilitators of and barriers to BARR implementation across all three cohorts.

[45] School attendance policies varied following COVID-19-related school closures and rendered much of the attendance data from March 2020 to the end of the school year unusable. For Cohort 3, the proportion of grade 9 students who missed 10 percent or more of enrolled school days between school start and February 28, 2020, was used for the analysis in cases where schools could not provide usable attendance data through the end of the school year.

[46] We also analyzed interview data from Cohort 1 control schools as well as interview data pertaining to COVID-19 from Cohort 3 control schools.

[47] Cohen's kappa coefficient is a statistical measure of interrater reliability that considers the amount of agreement that could be expected to occur through chance.

Measurement of Implementation Fidelity

We worked with the BARR Center to establish thresholds for each of the evaluated implementation components, defining the level at which BARR Center considered the BARR component to be adequately implemented at a school. Exhibit A4.1 provides information about the number of indicators for each key component on the structural component review form and the thresholds set to measure adequacy of implementation.

Exhibit A4.1. Number of Indicators and Thresholds Used to Calculate Fidelity

BARR Key Components	Indicators	Threshold
Professional Development	3	4.7 or higher
Restructuring Ninth Grade	7	5.3 or higher
Whole-Student Emphasis	4	5.5 or higher
Block/Team Meetings	8	5.6 or higher
I-Time	8	5.3 or higher
Risk Review	9	4.8 or higher
Contextual Support	7	4.0 or higher
Family Involvement	5	3.3 or higher

To calculate the component score for professional development, restructuring ninth grade, whole-student emphasis, contextual support, and family involvement, we averaged the BARR coordinator indicator ratings for each component. The BARR coordinator component score for each school was then compared against the predetermined threshold for adequate implementation. If the BARR coordinator component score matched or exceeded the threshold, school-level implementation of the component was deemed adequate. If it did not meet or exceed the threshold, school-level implementation of the component was deemed inadequate.

For the three observed BARR activities—block/team meetings, I-Time lessons, and risk review meetings—the final component fidelity rating was based on an equal weight of the BARR coordinator interview component score (50 percent) and observation component score (50 percent). To calculate the component score for block meetings, I-Time, and risk review, we employed the following method:

Following each observation, the AIR team members who observed BARR activities reviewed their observation ratings together and generated a consensus score for each component examined on the observation rubric. If multiple activities of the same type (e.g., block/team meeting) were observed in a school, the observation score was calculated by averaging all the observation ratings for a given activity. The observation score was then averaged with the BARR coordinator interview component score to create an overall component score for block/team

meetings, I-Time lessons, and risk review. If the overall component score matched or exceeded the threshold, school-level implementation of the component was deemed adequate. If it did not meet or exceed the threshold, school-level implementation for the component was deemed inadequate.

To meet adequate school-level implementation, schools were expected to meet or exceed four of the five priority component thresholds (restructuring, professional development, I-Time, block meeting, and risk review) and two of the three additional component thresholds (family involvement, whole-student emphasis, and contextual supports). To meet adequate cohort-level implementation, two thirds (67 percent) of a cohort of schools needed to have adequate implementation at the school level.

Measurement of BARR Impact on Teachers and Students

This section provides details on how survey measures were scaled and describes the models used to measure BARR impact on teachers' perception of self, students, and schools, and ninth-grade students' academic outcomes and behavior.

Scaling of Ninth-Grade Teachers' Survey Data

Exhibit A4.2 provides an overview of the reliability of the teacher survey scales used in the study. The survey items were scaled using the Rasch model for ordered response categories[48,49,50] to determine whether the items reliably measure the constructs they are intended to measure. Two reliability statistics were generated during the scaling process: the Rasch person separation reliability index (also referred to as Rasch reliability) and Cronbach's alpha statistic. Rasch reliability is a measure of how well the scale can distinguish among individuals of varying levels. Cronbach's alpha is a measure of the internal consistency of a scale. Internal consistency describes the extent to which all items in a scale measure the same concept. Reliability values for the two statistics range from 0 to 1, with values closest to 1 considered best and values of 0.7 or higher considered strong.

The reliability of seven of the eight teacher scales met or exceeded acceptable internal consistency expectations (Cronbach's alpha ranged from 0.71 to 0.87). The reliability of the postsecondary educational expectations scale was lower, with a Cronbach's alpha of .59.

[48] Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561–573.
[49] Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (Exp. ed.). University of Chicago Press.
[50] Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement.* MESA Press.

Exhibit A4.2. Teacher Survey Constructs and Reliability

Construct	Items (*n*)	*N*	Cronbach's Alpha
Strength-Based Mindset	5	512	0.75
Postsecondary Educational Expectations	4	509	0.59
Student Academic Motivation	6	506	0.87
Perception of Students' Behavior, Commitment, and Attitudes	8	505	0.86
View of Students' Observed Behavior, Commitment, and Attitudes	7	503	0.82
Use of Data	7	500	0.71
Teacher Self-Efficacy	8	496	0.83
Collaboration With and View of Colleagues	8	495	0.84
School Supports	4	495	0.82
Comfort With Technology*	8	158	0.90
Self-Efficacy for Instructional Use of Technology*	4	158	0.87

* This construct was only administered to Cohort 3 teachers.

Impact Analysis

To preserve the integrity of random assignment, all outcome analyses include all randomly assigned schools, even if they did not implement BARR after being assigned to the treatment group. (This happened in three schools in Cohort 1 and in one school in Cohort 2.) This is known as an intent-to-treat (ITT) design. ITT estimates capture the effect of having been *offered* BARR, regardless of whether the offer is then taken up. This ITT approach also extends to the student level. Students enrolled in study schools at the start of the school year are included in the ITT analysis. In eight treatment schools across the three cohorts, a subset of ninth-grade students was not included in the BARR model implementation for various logistical and scheduling reasons. The ITT estimates include these students as part of the treatment group.

We estimate program impacts by comparing outcomes for students in schools assigned to the treatment group to outcomes of students in schools assigned to the control group using the following general estimation model:

$$y_{is} = \beta_0 + \beta_1 Treat_s + x1'_{is}\beta_2 + x2'_s\beta_3 + x3'_s\beta_4 + \varepsilon_{is} \quad (1)$$

The variables in the model are as follows:

- y_{is} represents the outcome measure for student *i* in school *s*,

- $Treat_s$ is a binary variable indicating whether school *s* is a treatment school,
- $x1'_{is}$ is a vector of student-level variables, one for each the following student characteristics:
  - Prior test score at eighth grade
  - Whether the student is female
  - Whether the student is a student of color
  - Whether the student has an IEP (individualized education program)
  - Whether the student is an English learner (EL)
  - Whether the student is eligible for FRPL,
- $x2'_s$ is the school's prior graduation rate, an important school-level academic background characteristic,
- $x3'_s$ is a vector of block dummy variables identifying the randomization groups, and
- ε_{is} is the error term.

The parameter β_1 is the estimate of the impact of assignment to BARR on the student or teacher outcome.

Blocking and Differing Assignment Ratios

We grouped the schools into 30 random assignment blocks that were usually based on the local areas in which the schools are located. Schools within each block were then randomized into either the treatment or control group; therefore, within each block, schools had the same probability of assignment. However, the random assignment ratios varied across blocks. To account for these different assignment probabilities, a vector of block dummy variables was included in the impact model.[51]

Clustering Within Schools

We originally planned to estimate these impact models using a cross-cohort, multilevel hierarchical linear modeling (HLM) framework with school random effects to account for the clustering of students within schools (Raudenbush & Bryck, 2002) and cohort fixed effects. However, given the variation in context and implementation conditions across cohorts (including the COVID-19 pandemic), we decided to estimate impacts for each cohort separately, and combine the effects across cohorts using meta-analytical methods (see below). This then resulted in relatively small school-level sample sizes for each impact analysis, especially

[51] The block dummy variables mirrored randomization grouping to the extent possible. However, block dummies were reassigned in two special cases: (1) single random assignment schools assigned to single-school randomization blocks (five in total) were grouped together across outcome analyses, and (2) randomization groups for which there was not at least one treatment school and one control school with outcomes data were reassigned a block dummy for that series of outcome analyses. Considerations of random assignment probability, region, locale, and school demographics were used to reassign schools a block dummy in these special cases.

considering the need to include blocking variables to account for differences in random assignment ratios across blocks within cohorts. As a result, we found that the HLM models frequently failed to converge, especially for student subgroups and outcomes with a skewed distribution of values (e.g., suspensions, course failure).

When data are clustered, outcomes may be correlated across observations, and statistical models need to account for this to obtain accurate tests of statistical significance. HLM models are one way to address the correlation of observations with clusters. Because we were unable to consistently estimate HLM models, we decided to use an alternative approach to account for clustering: a more traditional linear regression model (i.e., without school random effects) with a correction to the standard errors of our impact estimates using the approach developed by Huber (1967) and White (1980). For one of our outcomes, course completion, we compared impact estimates from the HLM models that we were able to run successfully with those from models with Huber-White standard errors and found no meaningful differences in the point estimates. Huber-White standard errors were somewhat smaller. We found that, on average, where models converged, the robust standard errors were about 35 percent smaller than those from the random effects model. We concluded that it would be preferable to consistently use the less computationally intensive Huber-White method for all the impact estimates in this report, but some estimates that are statistically significant using this approach may not have been if we had been able to estimate them with a random-effects HLM model. See Maas & Hox (2004) and Abe & Gee (2014) for more background on the trade-offs between these analytical methods.

Effect Size Calculations

For continuous outcomes, we calculated effect sizes using the Hedge's g formula, which divides the estimated average treatment effect ($\hat{\beta}_1$) by the pooled standard deviation and applies a small sample size correction (What Works Clearinghouse, 2020). For dichotomous outcomes, we calculated effect sizes using the Cox index with regression-adjusted proportions (What Works Clearinghouse 2020).

Baseline Achievement

To account for differences in student baseline achievement across schools and research groups, the analytic models included, where available, student eighth-grade academic performance on state standardized tests in ELA and mathematics. To ensure that eighth-grade test scores across assessment and state were comparable, the analytic models included z-scores of students' eighth-grade test scores standardized within randomization block. When each school in a randomization block provided student-level scores for the same assessment, scores were simply standardized within the group to have a mean of 0 and a standard deviation of 1. However, when a randomization group had a mix of scores from different assessments and states, student eighth-

grade test scores were first translated into state assessment means and standard deviations to a common scale using state means and standard deviations from the National Assessment of Educational Progress (NAEP; U.S. Department of Education, 2017, 2019), and then standardized within randomization block (i.e., to mean 0 and a standard deviation of 1).[52]

There were 12 schools for which we did not have usable eighth-grade test score data. This is discussed further below in the section on missing data. To account for these gaps in baseline achievement differences, we also included the school-level, adjusted cohort graduation rate for the year prior to the study year as a second baseline measure in the analytic models.[53] We were able to obtain 64 out of 66 schools' adjusted cohort graduation rate data from EDFacts (U.S. Department of Education, n.d.).

Missing Data

Missing outcomes. Students with missing outcome data were excluded from the analytic sample for analyses involving those missing outcomes. For some models, all students in either the treatment or control condition from a random assignment block were missing outcome data. Reasons for this included (1) a school or schools had not provided data for a particular outcome, or (2) a school or schools had no students from that student sample with outcome data (i.e., having no ELs with data on credits earned). When there was not at least one student in the block with outcome data for either the treatment or control condition, we excluded the entire block from analysis.

Missing covariates. To avoid further reduction of the sample and maintain statistical power to detect effects, students with missing covariates were included in the analytic sample. We imputed missing student-level covariates with a 0 and created a corresponding missing indicator variable to include in the analysis models. Missing school-level data (e.g., percent eligible for FRPL) were imputed with the average value of that student-level variable (e.g., an indicator for the student being eligible for FRPL) in the sample, and no missing indicators were created for school-level variables.

For missing eighth-grade test scores, we took additional steps when a whole school was missing data. When students were missing eighth-grade test scores individually (i.e., when scores were available for some students but not all), we imputed their eighth-grade score to be the average score from their school and assigned a missing indicator. However, there were four study schools where we could not secure student eighth-grade test scores, and eight schools where we could not obtain usable assessment means with which to standardize scores. When an entire study

[52] NAEP composite scores are only available for the 2017 and 2019 NAEP administrations, corresponding to the 2016/17 and 2018/19 school years, respectively. To translate Cohort 2's eighth-grade scores from the 2017/18 school year onto the NAEP scale, we used a straight average of the 2017 and 2019 scores.

[53] Publicly available adjusted cohort graduation rates were obtained from EdFACTS.

school did not have usable eighth-grade test score data, we excluded (i.e., set to missing) the entire randomization block's eighth-grade data from the analytic model, given the need to standardize within group. For such blocks, we imputed values to 0 (the mean to which test scores in each randomization groups were standardized) for all students and schools in the group.[54]

All baseline covariates had some degree of missing data, with the eighth-grade academic performance on state standardized tests in ELA and mathematics and the FRPL status variable having the highest degree of missingness: 42.0 percent of students were missing the eighth-grade ELA score from the state test, 43.3 percent were missing the math score from the state test, 20.7 percent were missing their FRPL status, and fewer than 5 percent were missing data for each of the other covariates.

Subgroup Analyses

To address the exploratory research questions about the impacts of BARR on subgroups of students, we limited the sample for the corresponding impact analysis to the specific subgroup (e.g., male or female students, students eligible for FRPL). We then dropped the corresponding student-level baseline variable from the model.

Complier Average Causal Estimates to Account for School- and Student-Level Crossover

To complement our main intent-to-treat analysis, we conducted a treatment-on-the-treated analysis using complier average causal effect (CACE) analysis that adjusts for the reality that some students whose school was assigned to BARR did not receive the BARR intervention. In our study, there were two reasons for why a student did not receive BARR: (1) the student's school decided not to implement BARR even though it was assigned to do so ("crossover schools"), and (2) the student's school only implemented BARR for a portion of the school's ninth-grade student body for various logistical and scheduling reasons.

CACE analyses were conducted using an instrumental variable two-stage least squares model. In the first stage, a regression model predicted exposure to the BARR model based on the same regressor variables used in the ITT impact model (equation 1 above). The variable denoting exposure to BARR was coded as 1 in schools assigned to the treatment group except in the crossover and partial-implementing schools noted above. In crossover schools, this variable was coded as "0," indicating that students were not exposed to the BARR model. Some partial-implementing schools provided exact lists of which students received BARR and which students did not. In those schools, the exposure to the BARR indicator was coded accordingly. In other partial-implementing schools, the school only provided the proportion of students who

[54] We did not assign the student-level missing data indicator to 1. We did this so that the missing indicator would capture any effect of student-level missingness, not school-level missingness. Student-level missingness may be correlated with student-level factors, such as absenteeism or mobility. A school-level missing indicator would be correlated with the block dummies.

received the intervention. In those schools, we randomly assigned a "1" to that reported proportion of the students (i.e., they received BARR) and a "0" to the remainder. The predicted values from the first-stage regression were retained and used in the second stage.

In the second stage, the same ITT regression model was re-estimated using the predictions from the first stage as the independent variable of interest instead of the original random assignment variable (which indicated "assignment to the BARR condition" as opposed to actual exposure to the BARR model). The estimated coefficient on this newly created variable is the reported CACE estimate.

How We Combined Estimates Across Cohorts

To obtain results across the three cohorts, we used a fixed-effects meta-analysis model as is used by the U.S. Department of Education's What Works Clearinghouse to estimate intervention effects across different studies (What Works Clearinghouse, 2020). This model assumes that the different cohorts present unique evaluation environments and that differences in the estimated impacts reflect variation in the actual program effects in these different environments as opposed to being solely due to measurement error. This approach results in a weighted average of the effects across the cohorts where estimates with greater precision (i.e., lower standard errors) are weighted more heavily. Because sample size is related to precision, cohorts with a larger sample tend to get weighted more heavily in this approach. The estimated pooled average treatment effect from the model was tested for significance using a z-test.

We used the meta-analysis model to obtain the "difference" estimates for the combined cohorts (e.g., for "Cohorts 1, 2, and 3"). Other values in the reporting tables for the combined cohorts (e.g., adjusted means) were obtained by averaging across the cohorts using the same precision weights (i.e., the weights based on the standard error of the regression estimates). Hence, every across-cohort average was obtained in the same manner using the weights from the meta-analysis model for the "difference." Because of the variation across cohorts in how estimates turn into effect sizes (because every cohort had a different standard error for the outcome), the averaging of effect sizes across cohorts using the same precision weights produced incongruent results (e.g., opposite signs for the combined cohort average "difference" and average "effect size"). To obtain congruent results across cohort effect sizes, we (1) obtained a "conversion factor" for each cohort (estimate/effect size), (2) averaged the conversion factors across cohorts using the precision weights, and (3) divided the meta-analysis estimate by the average conversion factor to get the across-cohort effect size.

Measurement of BARR Scale-Up and Sustainability

We used the qualitative analysis software NVivo to code the transcripts and interview notes from BARR regional managers and BARR coaches for each year of data collection. We coded these interview data using a coding schema containing *a priori* codes based on the six strategies of BARR scale-up. During the coding process, we also added emergent codes to the coding structure to capture the variety of responses from BARR regional managers and BARR coaches.

To measure the extent to which BARR scale-up strategies were carried out as planned, we used data provided by BARR Center and Hazelden Publishing to calculate indicator scores for each BARR scale-up indicator. The remainder of this section describes how the scale-up indicator ratings were established and analyzed.

We assessed scale-up of the BARR model in 4 of the 5 years of the grant period (May 2017–May 2021). Measurement was based on predetermined thresholds set for each scale-up indicator. All thresholds for this analysis were established in consultation with BARR Center and based on theoretical expectations and practical experiences prior to and during the first 2 years of the study. Most thresholds were values on a continuous scale (e.g., 85 or more on a scale from 0 to 100). For example, the expectation set for the indicator "Schools Use BARR Communication Strategies" under the strategy "Expand School and District Awareness" was that 51 percent of schools that have implemented BARR for 1 year would need to use BARR communication strategies to meet or exceed BARR's expectations for model scale-up. A few of these indicators were binary (e.g., yes/no).

In cases where the scale-up threshold was based on a continuous scale (e.g., 0–100) and BARR Center met or exceeded the indicator threshold, we rated the indicator as "met." In cases where BARR Center demonstrated progress towards a continuous indicator but did not meet or exceed the threshold, we rated the indicator as "partially met." When an indicator threshold was binary (e.g., yes/no) and BARR met the threshold, we rated the indicator as "met." In cases where an indicator threshold was binary and not met, we rated the indicator as "not met." Appendix G provides details on the scale-up strategies and thresholds as well as detailed results for each year of the study.

A5. Characteristics of Study Schools and Summary of PSAT/NMSQT Testing Status by School and Cohort

Exhibit A5.1. Characteristics of Study Schools

Cohort	School	Locale	Total Enrollment	Students of Color (%)	English Learners (%)	Eligible for FRPL (%)	Special Education (%)	Prior-Year Graduation Rate
Cohort 1	A1	Rural	385	3.9	*	68.8	24.7	77.0
	A2	Rural	321	4.4	*	57.6	14.5	97.0
	A3	Suburb	748	23.3	*	54.1	17.3	77.0
	A4	City	480	98.3	*	100.0	31.9	54.5
	A5	City	306	*	*	*	*	54.5
	A6	Rural	232	3.4	0.0	43.5	15.9	94.5
	A7	Rural	996	2.1	0.0	58.9	18.0	84.5
	A8	City	1,365	39.7	3.7	57.0	17.6	84.5
	A9	Suburb	2,827	67.3	6.7	58.9	9.7	96.0
	A10	Town	1,800	88.4	22.7	92.9	12.7	90.0
	A11	Suburb	788	11.7	0.0	36.9	8.3	82.0
	A12	Suburb	1,689	78.9	5.1	52.0	10.2	95.0
	A13	City	802	24.4	0.3	53.1	21.1	80.0
	A14	Town	1,033	8.2	*	55.1	17.2	82.0
	A15	Rural	476	5.0	*	28.6	17.1	82.0
	A16	City	170	100.0	*	*	*	54.5
	A17	Rural	1,319	19.0	2.3	44.7	14.3	94.5
	A18	Suburb	2,624	88.5	12.1	78.1	15.3	95.0
	A19	Rural	1,847	95.6	12.9	85.5	17.9	88.0
	A20	City	1,573	95.7	17.5	92.6	16.0	92.0
	A21	City	2,370	90.9	9.6	82.2	8.8	80.0
	A22	Suburb	822	14.1	0.0	57.2	11.6	82.0
Cohort 2	B1	Suburb	559	90.7	17.0	79.1	15.9	87.0
	B2	City	1,543	52.5	26.8	54.7	14.5	80.0
	B3	Suburb	1,717	75.6	10.6	61.9	15.8	86.0
	B4	Rural	400	5.5	*	48.8	16.5	77.0
	B5	Suburb	1,228	97.0	32.0	84.1	20.0	58.0
	B6	Town	1,622	9.2	0.3	11.5	11.9	97.0
	B7	City	2,207	94.2	48.4	69.9	7.2	82.0
	B8	City	2,895	97.9	14.0	88.5	11.2	82.0
	B9	City	360	*	0.5	*	19.6	97.0
	B10	City	630	79.2	12.6	*	41.7	74.5

(continues)

Exhibit A5.1. Characteristics of Study Schools (continued)

Cohort	School	Locale	Total Enrollment	Students of Color (%)	English Learners (%)	Eligible for FRPL (%)	Special Education (%)	Prior-Year Graduation Rate
Cohort 2 (cont.)	B11	City	462	97.6	10.5	*	14.5	94.5
	B12	City	1,468	56.4	26.7	58.2	18.2	77.0
	B13	Rural	1,090	88.9	12.8	82.3	13.0	84.0
	B14	City	1,969	44.7	3.8	33.8	11.5	88.0
	B15	Suburb	1,841	40.1	6.6	*	17.9	82.0
	B16	Suburb	1,400	97.5	31.0	78.5	16.0	65.0
	B17	City	1,577	59.4	12.3	60.9	16.6	77.0
	B18	City	1,827	97.8	57.0	92.2	13.2	77.0
	B19	City	1,620	99.1	40.9	93.6	6.9	88.0
	B20	City	640	*	31.0	*	20.6	*
	B21	City	383	96.3	16.2	*	16.7	94.5
	B22	City	403	98.5	14.0	91.1	26.0	*
Cohort 3	C1	Suburb	1,262	12.9	0.3	43.0	*	88.0
	C2	City	1,915	69.7	28.2	52.3	9.5	92.0
	C3	City	1,597	57.9	6.4	26.0	4.7	96.0
	C4	Suburb	1,672	65.2	10.7	55.7	14.7	82.0
	C5	Rural	734	6.3	*	35.6	20.7	82.0
	C6	City	2,453	53.0	2.9	53.0	18.4	82.0
	C7	Suburb	2,424	98.5	21.0	91.3	8.7	92.0
	C8	Suburb	1,539	14.7	1.5	8.6	9.6	95.0
	C9	City	529	98.3	29.0	90.9	26.0	67.0
	C10	City	488	92.0	29.9	*	24.8	72.0
	C11	Suburb	1,591	85.9	12.8	*	17.9	92.0
	C12	Suburb	1,826	27.9	7.0	34.8	13.0	87.0
	C13	City	1,099	60.9	15.4	*	16.1	62.0
	C14	City	1,852	64.6	2.4	9.2	8.7	98.0
	C15	Suburb	1,683	60.5	9.3	50.5	14.6	87.0
	C16	Suburb	805	8.3	*	20.0	16.3	83.0
	C17	Rural	312	31.4	1.6	33.3	10.9	94.5
	C18	City	1,858	81.1	5.9	72.0	10.4	98.0
	C19	Suburb	1,672	98.9	27.0	67.7	7.0	83.0
	C20	City	354	94.9	5.0	87.3	20.0	96.0
	C21	City	535	95.1	48.4	*	23.6	57.0
	C22	City	853	98.8	*	68.6	17.5	84.0

* Indicates data are not available.

Source: Common Core of Data 2017–18. 2018–19, and 2019–20 used for school characteristics.

Exhibit A5.2. Summary of PSAT/NMSQT Testing Status by School and Cohort

Cohort	School	BARR/ Control	*N* Assigned	*N* Available to Be Tested	*N* Tested	Test Rate (%)	Test Rate Among Offered (%)
Cohort 1	A1	Treatment	76	59	54	71%	91%
	A2	Treatment	65	57	46	71%	81%
	A3	Treatment	165	147	132	80%	90%
	A4	Treatment	156	127	33	21%	26%
	A6	Treatment	63	61	57	90%	94%
	A7	Treatment	227	205	187	82%	91%
	A8	Treatment	301	248	223	74%	90%
	A9	Treatment	735	591	579	79%	98%
	A10	Treatment	464	319	287	62%	90%
	A11	Treatment	163	135	124	76%	92%
	A12	Treatment	416	363	341	82%	94%
	A13	Treatment	232	183	130	56%	71%
	A14	Control	269	235	202	75%	86%
	A15	Control	112	99	93	83%	94%
	A16	Control	19	26	10	53%	38%
	A17	Control	294	280	227	77%	81%
	A18	Control	719	319	246	34%	77%
	A20	Control	444	355	305	69%	86%
	A22	Control	202	194	167	83%	86%
Cohort 2	B1	Treatment	144	101	98	68%	97%
	B2	Treatment	381	311	221	58%	71%
	B3	Treatment	533	430	361	68%	84%
	B4	Treatment	97	84	71	73%	85%
	B5	Treatment	348	257	172	49%	67%
	B6	Treatment	395	360	338	86%	94%
	B7	Treatment	614	352	345	56%	98%
	B9	Treatment	76	59	44	58%	75%
	B10	Treatment	63	45	43	68%	95%
	B11	Treatment	101	88	84	83%	95%
	B12	Control	391	286	257	66%	90%
	B13	Control	357	236	201	56%	85%
	B14	Control	500	481	308	62%	64%
	B15	Control	566	272	220	39%	81%
	B16	Control	413	272	204	49%	75%
	B17	Control	358	272	204	57%	75%
	B18	Control	552	*	249	45%	*
	B19	Control	515	381	320	62%	84%
	B21	Control	104	94	89	86%	95%

* Indicates data are not available.

Source: American Institutes for Research calculations from school-provided PSAT data.

A6. Baseline Equivalence for Study Schools in the Administrative Data Sample

Exhibit A6. Baseline Equivalence for Study Schools in the Administrative Data Sample

	Treatment		Control		Standardized Difference	*P*-value
	Mean	SD	Mean	SD		
Cohort 1						
Graduation rate	85.7	0.10	90.2	0.06	-0.48	0.243
School characteristic						
Percent female	48.5	0.03	47.6	0.04	0.26	0.592
Percent students of color	57.9	0.28	58.5	0.40	-0.02	0.974
Percent FRPL	62.4	0.19	50.2	0.22	0.58	0.318
Percent EL	7.2	0.07	8.3	0.07	-0.16	0.746
Percent special education	15.7	0.06	17.0	0.02	-0.25	0.492
Total schools	15		7			
Total students	4,060		1,943			
Cohort 2						
Graduation rate	82.0	0.10	80.6	0.07	0.16	0.731
School characteristic						
Percent female	47.6	0.02	48.1	0.04	-0.14	0.753
Percent students of color	81.7	0.30	75.3	0.23	0.23	0.621
Percent FRPL	68.4	0.23	74.5	0.19	-0.28	0.555
Percent EL	23.3	0.19	28.1	0.22	-0.23	0.620
Percent special education	14.7	0.06	15.9	0.05	-0.21	0.654
Total schools	10		10			
Total students	3,332		3,490			
Cohort 3						
Graduation rate	88.2	0.06	87.4	0.12	0.09	0.845
School characteristic						
Percent female	48.8	0.03	47.7	0.02	0.37	0.412
Percent students of color	61.6	0.28	61.4	0.28	0.01	0.989
Percent FRPL	69.5	0.22	47.0	0.26	0.88	0.084
Percent EL	13.7	0.10	7.9	0.10	0.54	0.235
Percent special education	15.4	0.05	14.3	0.07	0.17	0.710
Total schools	11		10			
Total students	4,098		2,605			

(continues)

Exhibit A6. Baseline Equivalence for Study Schools in the Administrative Data Sample (continued)

	Treatment		Control		Standardized Difference	*P*-value
	Mean	SD	Mean	SD		
Cohorts 1, 2, and 3						
Graduation rate	85.5	0.09	85.1	0.10	0.04	0.873
School characteristic						
Percent female	48.3	0.03	47.8	0.03	0.15	0.563
Percent students of color	66.2	0.30	66.8	0.30	-0.02	0.933
Percent FRPL	66.6	0.21	60.4	0.26	0.26	0.360
Percent EL	14.4	0.14	16.9	0.19	-0.15	0.578
Percent special education	15.3	0.06	15.6	0.05	-0.07	0.789
Total schools	36		27			
Total students	11,490		8,038			

Note: SD = standard deviation; FRPL = free or reduced-price lunch; EL = English learner.

Source: Graduation rates obtained from the U.S. Department of Education's EDFacts and American Institutes for Research administrative data. The numbers are based on the Credits Earned variable and include students in the earlier joiner sample frame.

A7. Baseline Equivalence for Study Schools in the PSAT Data Sample (for All Schools for Which We Have PSAT-NMSQT Data)

Exhibit A7. Baseline Equivalence for Study Schools in the PSAT Data Sample

	Treatment		Control		Standardized Difference	*P*-value
	Mean	SD	Mean	SD		
Cohort 1						
Graduation rate	87.4	0.10	90.2	0.06	-0.30	0.491
School characteristic						
Percent female	48.7	0.04	47.6	0.04	0.27	0.578
Percent students of color	51.4	0.30	58.5	0.40	-0.20	0.710
Percent FRPL	54.3	0.10	50.2	0.22	0.25	0.715
Percent EL	6.8	0.08	8.3	0.07	-0.19	0.695
Percent special education	14.6	0.07	17.0	0.02	-0.42	0.314
Total schools	12		7			
Total students	2,193		1,250			
Cohort 2						
Graduation rate	82.4	0.11	80.6	0.07	0.19	0.688
School characteristic						
Percent female	48.1	0.03	48.1	0.04	0.01	0.987
Percent students of color	77.1	0.32	74.6	0.23	0.08	0.859
Percent FRPL	62.2	0.22	73.9	0.19	-0.54	0.271
Percent EL	25.6	0.22	28.6	0.22	-0.13	0.777
Percent special education	15.8	0.07	15.7	0.05	0.01	0.980
Total schools	10		9			
Total students	1,777		2,052			
Cohorts 1 and 2						
Graduation rate	85.0	0.11	84.0	0.08	0.10	0.745
School characteristic						
Percent female	48.4	0.04	47.9	0.04	0.13	0.702
Percent students of color	63.5	0.34	68.9	0.31	-0.16	0.625
Percent FRPL	58.5	0.18	67.2	0.23	-0.43	0.256
Percent EL	16.4	0.19	21.4	0.20	-0.25	0.459
Percent special education	15.2	0.07	16.2	0.04	-0.16	0.619
Total schools	22		16			
Total students	3,970		3,302			

Note: SD = standard deviation; FRPL = free or reduced-price lunch; EL = English learner.

Source: American Institutes for Research administrative data. The numbers are based on the PSAT Total Score variable and include students in the earlier joiner sample frame.

A8. Baseline Equivalence for Teacher Survey Data Sample (for All Schools for Which We Have Teacher Survey Data)

Exhibit A8. Baseline Equivalence for Study Schools in the Teacher Survey Data Sample

	Treatment		Control		Standardized Difference	*P*-value
	Mean	SD	Mean	SD		
Cohort 1						
Teacher characteristic						
Female	60.2	0.22	68.7	0.14	-0.41	0.320
Person of color	25.3	0.28	19.4	0.22	0.22	0.624
1 year experience	7.2	0.12	6.0	0.07	0.12	0.770
2–5 years' experience	31.3	0.20	26.9	0.17	0.23	0.616
6–10 years' experience	16.9	0.17	14.9	0.16	0.11	0.812
11+ years' experience	39.8	0.23	47.8	0.20	-0.35	0.449
Advanced degree	59.0	0.20	55.2	0.24	0.17	0.739
School characteristic						
Percent students of color	55.6	0.35	38.8	0.38	0.44	0.387
Percent FRPL	66.1	0.16	58.3	0.18	0.45	0.419
Percent EL	6.7	0.05	8.5	0.07	-0.28	0.727
School size	1,488	889	1,362	776	0.14	0.759
Percent rural	24.1	0.43	28.4	0.45	-0.09	0.849
Total schools	14		7			
Total teachers	81		66			
Cohort 2						
Teacher characteristic						
Female	48.2	0.11	54.6	0.15	-0.46	0.300
Person of color	16.4	0.12	32.4	0.27	-0.73	0.109
1 year experience	7.3	0.11	12.0	0.12	-0.40	0.374
2–5 years' experience	25.5	0.15	28.7	0.22	-0.16	0.712
6–10 years' experience	17.3	0.12	13.9	0.08	0.33	0.468
11+ years' experience	42.7	0.21	39.8	0.23	0.13	0.776
Advanced degree	59.1	0.28	67.6	0.22	-0.33	0.471
School characteristic						
Percent students of color	63.4	0.34	73.4	0.25	-0.33	0.502
Percent FRPL	55.3	0.24	70.6	0.23	-0.62	0.263
Percent EL	18.6	0.13	18.7	0.15	-0.01	0.987
School size	1,383	531	1,467	494	-0.15	0.744
Percent rural	5.8	0.23	13.9	0.35	-0.26	0.579
Total schools	10		11			
Total teachers	105		101			

(continues)

Exhibit A8. Baseline Equivalence for Study Schools in the Teacher Survey Data Sample (continued)

	Treatment		Control		Standardized Difference	*P*-value
	Mean	SD	Mean	SD		
Cohort 1						
Teacher characteristic						
Female	61.5	0.20	52.2	0.20	0.44	0.398
Person of color	19.8	0.20	17.9	0.25	0.08	0.879
1 year experience	0.0	0.00	0.0	0.00	0.00	NA
2–5 years' experience	27.1	0.18	20.9	0.16	0.34	0.500
6–10 years' experience	13.5	0.09	16.4	0.12	-0.26	0.631
11+ years' experience	53.1	0.19	56.7	0.21	-0.17	0.745
Advanced degree	74.0	0.13	65.7	0.21	0.47	0.405
School characteristic						
Percent students of color	58.3	0.30	46.8	0.30	0.37	0.476
Percent FRPL	50.2	0.23	44.0	0.20	0.26	0.632
Percent EL	11.0	0.10	7.2	0.04	0.42	0.361
School size	1,810	563	1,349	516	0.80	0.125
Percent rural	4.2	0.20	6.0	0.24	-0.08	0.881
Total schools	10		7			
Total teachers	93		66			
Cohorts 1, 2, and 3						
Teacher characteristic						
Female	56.1	0.19	57.9	0.18	-0.10	0.713
Person of color	20.1	0.21	24.8	0.26	-0.20	0.463
1 year experience	4.8	0.10	7.0	0.10	-0.22	0.419
2–5 years' experience	27.7	0.18	26	0.19	0.09	0.744
6–10 years' experience	15.9	0.13	14.9	0.12	0.08	0.751
11+ years' experience	45.3	0.22	46.7	0.23	-0.06	0.820
Advanced degree	64.0	0.23	63.6	0.23	0.02	0.951
School characteristic						
Percent students of color	59.4	0.33	56.0	0.34	0.10	0.711
Percent FRPL	57.1	0.23	59.2	0.23	-0.09	0.772
Percent EL	13.1	0.11	13.9	0.13	-0.07	0.830
School size	1,559	691	1,403	597	0.24	0.374
Percent rural	10.6	0.31	15.7	0.36	-0.15	0.588
Total schools	34		25			
Total teachers	279		233			

Note: SD = standard deviation; FRPL = free or reduced-price lunch; EL = English learner.

Source: American Institutes for Research survey data are used for teacher characteristics. Common Core of Data 2017–18, 2018–19, and 2019–20 are used for school characteristics.

Appendix B. BARR Regional Manager, BARR Coach, Teacher, and Administrator Interview Protocols

B1. BARR Regional Manager and Coach Interview Protocols (Scale-Up Measurement)

B1.1. Cohorts 1 and 2 BARR Coach Interview

Hello, I am (insert name). As you may recall, our organization, the American Institutes for Research (AIR), is leading a study of the BARR program, which seeks to understand the impact of the BARR model on students and teachers in schools across the country.

The interview will take up to 90 minutes and consists of open-ended questions about school implementation of BARR as well as scale-up and sustainability of the program. Your participation is voluntary, and you can stop at any time. We want to assure you these conversations will be used solely for research purposes. Given the small number of BARR regional managers and coaches, it may be possible for others to identify you even though we will not attribute any of your comments to you. To minimize identification, we will screen data for identifiable information and not reference you directly in reports.

We would like to record this discussion to make sure we do not miss anything in our conversation. This recording will only be used to verify the notes taken during the discussion and will not be shared with anyone outside of the research team. Is it okay if we record this interview? *[Note to interviewer: If yes, begin recording and ask participant again if it is okay to record so that we have their responses recorded. If no, please do your best to take detailed notes.]*

Do you agree to participate in this interview? *[Note to interviewer: If yes, continue with interview. If no, thank participant for time and end interview.]*

Do you have any questions before we begin?

BARR Implementation

9. How many schools did you coach in the [SCHOOL YEAR] school year?
    a. How many of these schools participated in the BARR Scale-Up Study?

10. What were your primary roles and responsibilities as a regional manager or coach?
    a. Probe: What types of support did you provide to BARR coordinators?
    b. Probe: What types of support did you provide to schools?

11. How would you characterize successful scale-up, or expansion, of the BARR model to more schools nationwide?

The next question is specifically asking about schools that started BARR during the [SCHOOL YEAR] school year.

12. Do you think the schools that began in the [SCHOOL YEAR] school year were ready to implement BARR?

13. What factors impacted school readiness to implement BARR?

The remaining questions apply to **any** BARR school you worked with during the [SCHOOL YEAR] school year. If applicable, please note **differences between schools participating in the BARR Scale-Up Study and BARR schools that are not participating in the study**. *[Note to interviewer: Prompt whether there were any differences between study and nonstudy schools if the interviewee does not make distinctions on their own.]*

14. Which of the eight BARR strategies were the most successfully implemented by your schools?
    a. Prompt: You mentioned [NUMBER] strategies. Did any of the other strategies stand out as successfully implemented?

15. Which of the eight BARR strategies were the most challenging to implement by your schools?
    a. Prompt: You mentioned [NUMBER] strategies. Did any of the other strategies stand out as a challenging?

16. How, if at all, did administrators and teachers in your schools adjust the eight BARR strategies to address the challenges they encountered?

17. How, if at all, did you adjust the supports you provided to schools to address the challenges they encountered?
    a. Probe: I-Time, risk review, block meetings

18. What additional resources and supports would help your schools implement BARR?

19. What opportunities were there for the BARR teams at your BARR schools to collaborate with BARR teams at other schools?
    a. Do you see any differences in implementation between schools that participated in these opportunities versus schools that did not?

School and District Awareness

20. How, if at all, have BARR schools expanded local awareness of the BARR program?
    a. Probes: Social media, events

21. Which of these activities have been most successful in promoting awareness of the BARR model to other schools and districts in your region(s)?

22. Have these efforts directly or indirectly led to adoption of the program by other schools or districts? Please provide examples.

23. What additional resources and supports could current BARR schools use to promote the BARR program to other schools and districts in your region(s)?

Sustainable Cost of Services

24. What barriers have your schools encountered when trying to secure funding to cover the financial cost of the BARR program?

25. What approaches or strategies have schools used to successfully secure funding for the BARR program?
    a. Probe: District, state, philanthropic or other third-party funding

26. What financial resources and supports would help your schools as they try to secure funding for the program?

School-Level Staff Training at Scale

27. What were the most common challenges when providing school-level staff training?

28. How did you address or overcome these school-level staff training challenges?

29. What resources or supports would help improve school-level staff training?

Coach Training at Scale

30. What aspects of coaching have been the most challenging to master?

31. How, if at all, could regional manager and coach training be improved to address these challenges?

32. What resources would help regional managers and coaches better support schools?

BARR Infrastructure and Sustainability

33. What strategies have you used to build BARR visibility in your region?
    a. Which strategies have been most successful in building program visibility?

34. What additional supports would help you promote program awareness in your region?

35. What, if any, infrastructure is needed to support the schools currently implementing BARR?
    a. Probes: Tools, staffing, facilities

36. What, if any, infrastructure would be needed if more schools choose to adopt the model?
    a. Probes: Tools, staffing, facilities

Additional Comments

37. Is there anything you would like to share about BARR that you didn't have the opportunity to share already?

B1.2. Cohort 3 BARR Coach Interview

Hello, I am (insert name). As you may recall, our organization, the American Institutes for Research (AIR), is leading a study of the BARR program, which seeks to understand the impact of the BARR model on students and teachers in schools across the country.

The interview will take up to 90 minutes and consists of open-ended questions about school implementation of BARR as well as scale-up and sustainability of the program. Your participation is voluntary, and you can stop at any time. We want to assure you these conversations will be used solely for research purposes. Given the small number of BARR regional managers and coaches, it may be possible for others to identify you even though we will not attribute any of your comments to you. To minimize identification, we will screen data for identifiable information and not reference you directly in reports.

We would like to record this discussion to make sure we do not miss anything in our conversation. This recording will only be used to verify the notes taken during the discussion and will not be shared with anyone outside of the research team. Is it okay if we record this interview? *[Note to interviewer: If yes, begin recording and ask participant again if it is okay to record so that we have their responses recorded. If no, please do your best to take detailed notes.]*

Do you agree to participate in this interview? *[Note to interviewer: If yes, continue with interview. If no, thank participant for time and end interview.]*

Do you have any questions before we begin?

BARR Implementation

1. How many schools did you coach in the [SCHOOL YEAR] school year?
   a. How many of these schools participated in the BARR Scale-Up Study?

2. Prior to school closures, what were your primary roles and responsibilities as a regional manager or coach?
   a. Probe: What types of support did you provide to BARR coordinators?
   b. Probe: What types of support did you provide to schools?

3. After school campus closures, did your roles and responsibilities as a regional manager or coach change in any way? If so, please explain.

4. Following school closures, did you start working with new schools or schools with a later start date (such as fall 2020)?
   a. If not answered, probe for names of control schools.

5. How would you characterize successful scale-up, or expansion, of the BARR model to more schools nationwide?

The next question is specifically asking about schools that started BARR during the [SCHOOL YEAR] school year.

6. Do you think the schools that began in the [SCHOOL YEAR] school year were ready to implement BARR?
    a. Probe: What factors impacted school readiness to implement BARR?

7. After school campus closures, what factors impacted a school's ability to continue implementing BARR?

The remaining questions apply to **any** BARR school you worked with during the [SCHOOL YEAR] school year. If applicable, please note **differences between schools participating in the BARR Scale-Up Study and BARR schools that are not participating in the study**. *[Note to interviewer: Prompt whether there were any differences between study and nonstudy schools if the interviewee does not make distinctions on their own.]*

8. Prior to school campus closures, which of the eight BARR strategies were the most successfully implemented by your schools?

9. Prior to school campus closures, which of the eight BARR strategies were the most challenging to implement by your schools?

10. After school campus closures, which of the eight BARR strategies were the most successfully implemented by your schools?
    a. Prompt: How, if at all, did administrators and teachers in your schools adjust these strategies in response to school campus closures?

11. After school campus closures, which of the eight BARR strategies were the most challenging to implement by your schools?
    a. Prompt: How, if at all, did administrators and teachers in your schools adjust these strategies?

12. How, if at all, did you adjust the supports you provided to schools to address the challenges they encountered?
    a. Probe: Prior to school campus closures, after school campus closures

13. What additional resources and supports would help your schools implement BARR?
    a. Probe: After school campus closures, did schools ask for any specific supports? If so, please explain.

14. What opportunities were there for the BARR teams at your BARR schools to collaborate with BARR teams at other schools?
    a. Do you see any differences in implementation between schools that participated in these opportunities versus schools that did not?
        i. Probe: Prior to school campus closures, after school campus closures

School and District Awareness

15. How, if at all, have BARR schools expanded local awareness of the BARR program?
    a. Probes: Social media, events

16. Which of these activities have been most successful in promoting awareness of the BARR model to other schools and districts in your region(s)?

17. What additional resources and supports could current BARR schools use to promote the BARR program to other schools and districts in your region(s)?

Sustainable Cost of Services

18. What barriers have your schools encountered when trying to secure funding to cover the financial cost of the BARR program?

19. What approaches or strategies have schools used to successfully secure funding for the BARR program?
    a. Probe: District, state, philanthropic or other third-party funding

20. What financial resources and supports would help your schools as they try to secure funding for the program?

School-Level Staff Training at Scale

21. What were the most common challenges when providing school-level staff training?

22. How did you address or overcome these school-level staff training challenges?

23. What resources or supports would help improve school-level staff training?

Coach Training at Scale

24. What aspects of coaching have been the most challenging to master?

25. How, if at all, could regional manager and coach training be improved to address these challenges?

26. What resources would help regional managers and coaches better support schools?

BARR Infrastructure and Sustainability

27. What strategies have you used to build BARR visibility in your region?
    a. Which strategies have been most successful in building program visibility?

28. What additional supports would help you promote program awareness in your region?

29. What, if any, infrastructure is needed to support the schools currently implementing BARR?
    a. Probes: Tools, staffing, facilities

30. What, if any, infrastructure would be needed if more schools choose to adopt the model?
    a. Probes: Tools, staffing, facilities

Additional Comments

31. Is there anything you would like to share about BARR that you didn't have the opportunity to share already?

B2. Teacher Interview Protocols

B2.1. Cohort 1 BARR Teacher Interview

Hello, my name is _____ and I work for the American Institutes for Research (AIR). I am here today as part of our work evaluating the implementation of the BARR initiative this year.

Please note, this is not an evaluation of you personally. We would simply like to hear about your experience and your opinions related to the implementation of BARR at your school.

This discussion should take no more than 30–45 minutes. Please take a moment to review the attached interview consent form for more detailed information about this study and your rights as a participant. Do you have any questions before we get started?

I would like to be able to record our interview in order to accurately capture everything you tell me. The recording is purely for AIR's purposes and will not be shared with anyone else. Do I have your permission to record this interview*?* [If **yes**, begin recording. If **no**, take detailed notes of responses.]

A. Opening Questions

1. As background, what subject(s) and grade level(s) do you teach?

2. How long have you been a teacher at this school?
   a. *How long have you been a teacher in total?*

3. Could you please describe your personal experience teaching ninth graders this year?
   a. In what ways, if any, do you think being a BARR teacher has had an impact on this experience?

B. Professional Development

4. What professional development opportunities have you participated in this past year?
   a. What BARR-specific professional development did you participate in? (i.e., Foundational training? Periodic trainings during the school year?)

5. To what extent do you feel this professional development changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
   a. Probe: If so, please provide an example of what has changed for you this year.

6. Were there any aspects of this professional development that were particularly helpful to you as a teacher this year? Please provide examples.

7. Were there any aspects of this professional development that you found confusing or difficult to integrate into your teaching?

C. Ninth-Grade Structure

8. Please describe the current ninth-grade structure at your school. (Prompts: Related to teacher teams, common planning times, class size, student leveling)
    a. Probe: In what ways, if any, was the ninth grade restructured this year due to the implementation of the BARR model?

9. To what extent do you feel this ninth-grade structure has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

D. Team Meetings

10. Please describe some of the practices you use (alone or with other teachers) to monitor student progress.
    a. In what ways, if any, have you used student-level data to guide your instructional decisions?
    b. How important is it for you to be able to access real-time student-level data during the day?

11. Are you currently part of a team structure at your school? Do you share the same students across core classes? Do you have departmental teams? (If on multiple teams, probe for each.)
    a. [If yes to teams:] Do you have regular team meetings?
    b. How often do you meet with your team?
    c. In what ways, if any, are you able to discuss student progress on your team?

12. To what extent do you feel having team meetings has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

13. Are there any particular aspects of the team meetings that you think helped you as a teacher this year? Please provide examples.
    a. Collaborating with other teachers? (Probe: Instructional decisions)
    b. Building positive peer-to-peer relationships? (Probe: Feel respect, trust)

14. Are there any aspects of team meetings that you found challenging or difficult this year? Please provide examples.

E. Whole Student Approach

15. In what ways, if any, do you think students express:
    a. Their academic needs and strengths in the classroom?
    b. Their nonacademic needs and strengths in the classroom (i.e., physically, emotionally, socially)?

16. Please describe some of the practices you have used this year:
    a. To address students' needs in the classroom, if any.
    b. To build on students' strengths in the classroom, if any.

17. To what extent do you feel the way you address students' needs and strengths has changed this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)

a. Probe: If so, please provide an example of what has changed for you this year.
b. If so, to what do you attribute this change this year?

F. Social-Emotional Learning/Youth Development

18. Did you facilitate any social-emotional learning (SEL) or youth development activities/lessons in your classroom this year? (i.e., I-Times, but also probe for any other SEL activities/lessons)
    a. If so, were there any particular activities or lessons that were helpful to you as a teacher this year? Please provide examples.

19. To what extent do you feel using social-emotional learning or youth development activities/lessons has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

20. In what ways, if any, do you think incorporating social-emotional learning or youth development activities/lessons have affected:
    a. Developing positive teacher-to-student relationships?
    b. Encouraging positive student-to-student relationships in your classroom?

G. Student Referral Process

21. Please describe your experience with the current process for referring high-need students for supports and services (e.g., risk review, Student Assistance Team).
    a. If you have referred students, what is the level of feedback or follow-up you have received from this process?

22. To what extent do you feel the current student referral process has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

H. Parent Involvement

23. Please describe some of the practices or activities you have used this year to try and foster positive parent-teacher relationships.
    a. Probe: Individually?
    b. As a team?

24. To what extent do you feel your practices or activities related to parent involvement have changed this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

25. To what extent have you seen actual parent involvement change this school year compared to last year? (Better, the same, worse)

I. Contextual Support

26. What types of support, if any, did the school administration provide to you that helped you this year? Please provide examples.
    a. Probe: What kinds of support, if any, were provided by school administration specifically for BARR implementation?

27. To what extent do you feel the support from school administration changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

28. What kinds of support, if any, were provided by other school staff (e.g., academic coach, department head, BARR coordinator) that helped you this year? Please provide examples.
    a. Do you think the BARR coordinator's role is an important one? Why or why not?

29. To what extent do you feel having support from other school staff (e.g., academic coach, department head, BARR coordinator) changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

J. Overall Experience

30. Compared to ninth-grade students you may have had in the past, what do you think the school experience has been like for the group you had this year?
    a. Probe: What do you think the levels of achievement, attendance, and behavior have been like for the group of students you had this year? (Better, the same, worse)

31. Were there any major external or internal challenges that impacted your school this year?
    a. Were there any major challenges or barriers specific to implementing BARR at your school this year? (Probe for any aspects of the BARR model that may not have been discussed yet.)

32. Thinking about your overall experience this year with BARR, is there something that stands out above everything else? Please explain.

33. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

B2.2. Cohort 1 Control Teacher Interview

Hello, my name is _____ and I work for the American Institutes for Research (AIR). I am here today as part of our work evaluating the nationwide implementation of a ninth-grade program, and your school served as one of our control schools this year.

Please note, this is not an evaluation of you personally. We would simply like to hear about your experience and your opinions related to different aspects of teaching at your school.

This discussion should take no more than 30–45 minutes. Please take a moment to review the attached interview consent form for more detailed information about this study and your rights as a participant. Do you have any questions before we get started?

I would like to be able to record our interview in order to accurately capture everything you tell me. The recording is purely for AIR's purposes and will not be shared with anyone else. Do I have your permission to record this interview*?* [If **yes**, begin recording. If **no**, take detailed notes of responses.]

A. Opening Questions

1. As background, what subject(s) and grade level(s) do you teach?

2. How long have you been a teacher at this school?
   a. How long have you been a teacher in total?

3. Could you please describe your personal experience teaching ninth graders this year?

B. Professional Development

4. What professional development opportunities have you participated in this past year?

5. To what extent do you feel this professional development changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
   a. Probe: If so, please provide an example of what has changed for you this year.

6. Were there any aspects of this professional development that were particularly helpful to you as a teacher this year? Please provide examples.

7. Were there any aspects of this professional development that you found confusing or difficult to integrate into your teaching?

C. Ninth-Grade Structure

8. Please describe the current ninth-grade structure at your school. (Prompts: Related to teacher teams, common planning times, class size, student leveling)
   a. Did any changes occur this year to the ninth grade?

9. To what extent do you feel this ninth-grade structure has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
   a. Probe: If so, please provide an example of what has changed for you this year.

D. Team Meetings

10. Please describe some of the practices you use (alone or with other teachers) to monitor student progress.
    a. In what ways, if any, have you used student-level data to guide your instructional decisions?
    b. How important is it for you to be able to access real-time student-level data during the day?

11. Are you currently part of a team structure at your school? Do you share the same students across core classes? Do you have departmental teams? (If on multiple teams, probe for each.)
    a. [If yes to teams:] Do you have regular team meetings?
    b. How often do you meet with your team?
    c. In what ways, if any, are you able to discuss student progress on your team?

12. To what extent do you feel having team meetings has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

13. Are there any particular aspects of team meetings that you think helped you as a teacher this year? Please provide examples.
    a. Collaborating with other teachers? (Probe: Instructional decisions)
    b. Building positive peer-to-peer relationships? (Probe: Feel respect, trust)

14. Are there any aspects of team meetings that you found challenging or difficult this year? Please provide examples.

E. Whole Student Approach

15. In what ways, if any, do you think students express:
    a. Their academic needs and strengths in the classroom?
    b. Their nonacademic needs and strengths in the classroom (i.e., physically, emotionally, socially)?

16. Please describe some of the practices you have used this year:
    a. To address students' needs in the classroom, if any.
    b. To build on students' strengths in the classroom, if any.

17. To what extent do you feel the way you address students' needs and strengths has changed this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.
    b. If so, to what do you attribute this change this year?

F. Social-Emotional Learning/Youth Development

18. Did you facilitate any social-emotional learning or youth development activities/lessons in your classroom this year?
    a. If so, were there any particular activities or lessons that were helpful to you as a teacher this year? Please provide examples.

19. To what extent do you feel using social-emotional learning or youth development activities/lessons has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

20. In what ways, if any, do you think incorporating social-emotional learning or youth development activities/lessons has affected:
    a. Developing positive teacher-to-student relationships?
    b. Encouraging positive student-to-student relationships in your classroom?

G. Student Referral Process

21. Please describe your experience with the current process for referring high-need students for supports and services (e.g., Student Assistance Team).
    b. If you have referred students, what is the level of feedback or follow-up you have received from this process?

22. To what extent do you feel the current student referral process has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

H. Parent Involvement

23. Please describe some of the practices or activities you have used this year to try and foster positive parent-teacher relationships.
    a. Probe: Individually?
    b. As a team?

24. To what extent do you feel your practices or activities related to parent involvement have changed this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

25. To what extent have you seen actual parent involvement change this school year compared to last year? (Better, the same, worse)

I. Contextual Support

26. What types of support, if any, did the school administration provide to you that helped you this year? Please provide examples.

27. To what extent do you feel the support from school administration changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

28. What kinds of support, if any, were provided by other school staff (e.g., academic coach, department head) that helped you this year? Please provide examples.

29. To what extent do you feel having support from other school staff (e.g., academic coach, department head) changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

J. Overall Experience

30. Compared to ninth-grade students you may have had in the past, what do you think the school experience has been like for the group you had this year?

    a. Probe: What do you think the levels of achievement, attendance, and behavior have been like for the group of students you had this year? (Better, the same, worse)

31. Were there any major external or internal challenges that impacted your school this year?

32. As a control school in this study, your school will likely be implementing the BARR (Building Assets, Reducing Risks) model in ninth grade next year. Have you heard any information about this program yet?
    a. If given the option, would you be interested in implementing BARR in your classroom? Why or why not?

33. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

B2.3. Cohort 2 BARR Teacher Interview

Hello, my name is _____ and I work for the American Institutes for Research (AIR). I am here today as part of our work evaluating the implementation of the BARR initiative this year.

Please note, this is not an evaluation of you personally. We would simply like to hear about your experience and your opinions related to the implementation of BARR at your school.

This discussion should take no more than the period we have allotted (30–45 minutes). Please take a moment to review the interview consent form for more detailed information about this study and your rights as a participant. Do you have any questions before we get started?

I would like to be able to record our interview in order to accurately capture everything you tell me. The recording is purely for AIR's purposes and will not be shared with anyone else. Do I have your permission to record this interview*?* [If **yes**, begin recording. If **no**, take detailed notes of responses.]

A. Opening Questions

1. As background, what subject(s) and grade level(s) do you teach?

2. How long have you been a teacher at this school?
   a. How long have you been a teacher in total?

3. Could you please describe your personal experience teaching ninth graders this year?
   a. In what ways, if any, do you think being a BARR teacher has had an impact on this experience?

B. Professional Development

4. What professional development opportunities have you participated in this past year?
   a. What BARR-specific professional development did you participate in? (i.e., Foundational training? Periodic trainings during the school year?)

5. To what extent do you feel this professional development changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
   a. Probe: If so, please provide an example of what has changed for you this year.

6. Were there any aspects of this professional development that were particularly helpful to you as a teacher this year? Please provide examples.

7. Were there any aspects of this professional development that you found confusing or difficult to integrate into your teaching?

C. Ninth-Grade Structure

8. Please describe the current ninth-grade structure at your school. (Prompts: Related to teacher teams, common planning times, class size, student leveling)
   a. Probe: In what ways, if any, was the ninth grade restructured this year due to the implementation of the BARR model?

9. To what extent do you feel this ninth-grade structure has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

D. Team Meetings

10. Please describe some of the practices you use (alone or with other teachers) to monitor student progress.
    a. In what ways, if any, have you used student-level data to guide your instructional decisions?
    b. How important is it for you to be able to access real-time student-level data during the day?

11. Are you currently part of a team structure at your school? Do you share the same students across core classes? Do you have departmental teams? (If on multiple teams, probe for each.)
    a. [If yes to teams:] Do you have regular team meetings?
    b. How often do you meet with your team?
    c. In what ways, if any, are you able to discuss student progress on your team?

12. To what extent do you feel having team meetings has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

13. Are there any particular aspects of the team meetings that you think helped you as a teacher this year? Please provide examples.
    a. Collaborating with other teachers? (Probe: Instructional decisions)
    b. Building positive peer-to-peer relationships? (Probe: Feel respect, trust)

14. Are there any aspects of team meetings that you found challenging or difficult this year? Please provide examples.

E. Whole Student Approach

15. In what ways, if any, do you think students express:
    a. Their academic needs and strengths in the classroom?
    b. Their nonacademic needs and strengths in the classroom (i.e., physically, emotionally, socially)?

16. Please describe some of the practices you have used this year:
    a. To address students' needs in the classroom, if any.
    b. To build on students' strengths in the classroom, if any.

17. To what extent do you feel the way you address students' needs and strengths has changed this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.
    b. If so, to what do you attribute this change this year?

F. Social-Emotional Learning/Youth Development

18. Did you facilitate any social-emotional learning (SEL) or youth development activities/lessons in your classroom this year? (i.e., I-Times, but also probe for any other SEL activities/lessons)
    a. If so, were there any particular activities or lessons that were helpful to you as a teacher this year? Please provide examples.

19. To what extent do you feel using social-emotional learning or youth development activities/lessons has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

20. In what ways, if any, do you think incorporating social-emotional learning or youth development activities/lessons has affected:
    a. Developing positive teacher-to-student relationships?
    b. Encouraging positive student-to-student relationships in your classroom?

G. Student Referral Process

21. Please describe your experience with the current process for referring high-need students for supports and services (e.g., risk review, Student Assistance Team).
    a. If you have referred students, what is the level of feedback or follow-up you have received from this process?

22. To what extent do you feel the current student referral process has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

H. Parent Involvement

23. Please describe some of the practices or activities you have used this year to try and foster positive parent-teacher relationships.
    a. Probe: Individually?
    b. As a team?

24. To what extent do you feel your practices or activities related to parent involvement have changed this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

25. To what extent have you seen actual parent involvement change this school year compared to last year? (Better, the same, worse)

I. Contextual Support

26. What types of support, if any, did the school administration (e.g., principal, assistant principals) provide to you that helped you this year? Please provide examples.
    a. Probe: What kinds of support, if any, were provided by school administration specifically for BARR implementation?

27. To what extent do you feel the support from school administration (e.g., principal, assistant principals) changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

28. What kinds of support, if any, were provided by other school staff (e.g., academic coach, department head, BARR coordinator, colleagues) that helped you this year? Please provide examples.
    a. Do you think the BARR coordinator's role is an important one? Why or why not?

29. To what extent do you feel having support from other school staff (e.g., academic coach, department head, BARR coordinator, colleagues) changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

J. Overall Experience

30. Compared to ninth-grade students you may have had in the past, what do you think the school experience has been like for the group you had this year?
    a. Probe: What do you think the levels of achievement, attendance, and behavior have been like for the group of students you had this year? (Better, the same, worse)

31. Were there any major external or internal challenges that impacted your school this year?
    a. Were there any major challenges or barriers specific to implementing BARR at your school this year? (Probe for any aspects of the BARR model that may not have been discussed yet.)

32. Thinking about your overall experience this year with BARR, is there something that stands out above everything else? Please explain.

33. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

B2.4. Cohort 2 Control Teacher Interview

Hello, my name is _____ and I work for the American Institutes for Research (AIR). I am here today as part of our work evaluating the nationwide implementation of a ninth-grade program, and your school served as one of our control schools this year.

Please note, this is not an evaluation of you personally. We would simply like to hear about your experience and your opinions related to different aspects of teaching at your school.

This discussion should take no more than the period we have allotted (30–45 minutes). Please take a moment to review the interview consent form for more detailed information about this study and your rights as a participant. Do you have any questions before we get started?

I would like to be able to record our interview in order to accurately capture everything you tell me. The recording is purely for AIR's purposes and will not be shared with anyone else. Do I have your permission to record this interview*?* [If **yes**, begin recording. If **no**, take detailed notes of responses.]

A. Opening Questions

1. As background, what subject(s) and grade level(s) do you teach?

2. How long have you been a teacher at this school?
   a. How long have you been a teacher in total?

3. Could you please describe your personal experience teaching ninth graders this year?

B. Professional Development

4. What professional development opportunities have you participated in this past year?

5. To what extent do you feel this professional development changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
   a. Probe: If so, please provide an example of what has changed for you this year.

6. Were there any aspects of this professional development that were particularly helpful to you as a teacher this year? Please provide examples.

7. Were there any aspects of this professional development that you found confusing or difficult to integrate into your teaching?

C. Ninth-Grade Structure

8. Please describe the current ninth-grade structure at your school. (Prompts: Related to teacher teams, common planning times, class size, student leveling)
   a. Did any changes occur this year to the ninth grade?

9. To what extent do you feel this ninth-grade structure has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
   a. Probe: If so, please provide an example of what has changed for you this year.

D. Team Meetings

10. Please describe some of the practices you use (alone or with other teachers) to monitor student progress.
    a. In what ways, if any, have you used student-level data to guide your instructional decisions?
    b. How important is it for you to be able to access real-time student-level data during the day?

11. Are you currently part of a team structure at your school? Do you share the same students across core classes? Do you have departmental teams? (If on multiple teams, probe for each.)
    a. [If yes to teams:] Do you have regular team meetings?
    b. How often do you meet with your team?
    c. In what ways, if any, are you able to discuss student progress on your team?

12. To what extent do you feel having team meetings has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

13. Are there any particular aspects of team meetings that you think helped you as a teacher this year? Please provide examples.
    a. Collaborating with other teachers? (Probe: Instructional decisions)
    b. Building positive peer-to-peer relationships? (Probe: Feel respect, trust)

14. Are there any aspects of team meetings that you found challenging or difficult this year? Please provide examples.

E. Whole Student Approach

15. In what ways, if any, do you think students express:
    a. Their academic needs and strengths in the classroom?
    b. Their nonacademic needs and strengths in the classroom (i.e., physically, emotionally, socially)?

16. Please describe some of the practices you have used this year:
    a. To address students' needs in the classroom, if any.
    b. To build on students' strengths in the classroom, if any.

17. To what extent do you feel the way you address students' needs and strengths has changed this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.
    b. If so, to what do you attribute this change this year?

F. Social-Emotional Learning/Youth Development

18. Did you facilitate any social-emotional learning or youth development activities/lessons in your classroom this year?
    a. If so, were there any particular activities or lessons that were helpful to you as a teacher this year? Please provide examples.

19. To what extent do you feel using social-emotional learning or youth development activities/lessons has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

20. In what ways, if any, do you think incorporating social-emotional learning or youth development activities/lessons has affected:
    a. Developing positive teacher-to-student relationships?
    b. Encouraging positive student-to-student relationships in your classroom?

G. Student Referral Process

21. Please describe your experience with the current process for referring high-need students for supports and services (e.g., Student Assistance Team).
    a. If you have referred students, what is the level of feedback or follow-up you have received from this process?

22. To what extent do you feel the current student referral process has changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

H. Parent Involvement

23. Please describe some of the practices or activities you have used this year to try and foster positive parent-teacher relationships.
    a. Probe: Individually?
    b. As a team?

24. To what extent do you feel your practices or activities related to parent involvement have changed this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

25. To what extent have you seen actual parent involvement change this school year compared to last year? (Better, the same, worse)

I. Contextual Support

26. What types of support, if any, did the school administration (e.g., principal, assistant principals) provide to you that helped you this year? Please provide examples.

27. To what extent do you feel the support from school administration (e.g., principal, assistant principals) changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

28. What kinds of support, if any, were provided by other school staff (e.g., academic coach, department head, colleagues) that helped you this year? Please provide examples.

29. To what extent do you feel having support from other school staff (e.g., academic coach, department head, colleagues) changed your teaching practice this year compared to last year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what has changed for you this year.

J. Overall Experience

30. Compared to ninth-grade students you may have had in the past, what do you think the school experience has been like for the group you had this year?
    a. Probe: What do you think the levels of achievement, attendance, and behavior have been like for the group of students you had this year? (Better, the same, worse)

31. Were there any major external or internal challenges that impacted your school this year?

32. As a control school in this study, your school will likely be implementing the BARR (Building Assets, Reducing Risks) model in ninth grade next year. Have you heard any information about this program yet?
    a. If given the option, would you be interested in implementing BARR in your classroom? Why or why not?

33. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

B2.5. Cohort 3 BARR Teacher Interview

Hello, my name is _____ and I work for the American Institutes for Research (AIR). I am speaking with you today as part of our work evaluating the implementation of the BARR initiative.

In general, we'll ask about your experiences for the entire 2019–20 school year. Please note, this is not an evaluation of you personally. We would simply like to hear about your experiences and opinions related to the implementation of BARR at your school during the 2019–20 school year.

This discussion should take no more than the hour we have allotted. Please take a moment to review the interview consent form if you have not done so already. There is more detailed information about this study and your rights as a participant. Please feel free to stop me at any time if you'd like to take a break, reschedule the conversation, or prefer to discontinue the conversation. You can always choose not to answer specific questions. Do you have any questions before we get started?

I would like to be able to record our interview in order to accurately capture everything you tell me. The recording is purely for AIR's purposes and will not be shared with anyone else. Do I have your permission to record this interview*?* [If **yes**, begin recording. If **no**, take detailed notes of responses.]

A. Opening Questions (2 minutes)

1. As background, what subject(s) and grade level(s) did you teach during the 2019–20 school year?

2. How long had you been a teacher at this school during the 2019–20 school year?
   a. How long had you been a teacher in total in the 2019–20 school year?

3. In what ways, if any, did you think being a BARR teacher had an impact on teaching ninth graders during the 2019–20 school year prior to the school campus closure due to COVID-19?
   a. After the school campus closure?

B. Professional Development (5 minutes)

4. What BARR-specific professional development did you participate in during the 2019–20 school year prior to the school campus closure? (i.e., Foundational training? Periodic trainings during the school year?)
   a. During the school campus closure, what BARR-specific trainings did you participate in, if any?

*For the next and subsequent questions throughout the interview, I will ask you to quantify your response by saying major change, moderate change, minor change, or no change.

5. To what extent do you feel BARR-specific professional development provided changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
   a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

6. Were there any aspects of BARR-specific professional development that were particularly helpful to you as a teacher during the 2019–20 school year? Please provide examples. (i.e., delivery method of professional development).

7. Were there any aspects of BARR-specific professional development that you found confusing or difficult to integrate into your teaching, or understand?

C. Ninth-Grade Structure (5 minutes)

8. Please describe the BARR ninth-grade structure at your school prior to the school campus closure during the 2019–20 school year. (Prompts: Related to teacher teams, class size, student leveling)
    a. Probe: In what ways, if any, was the ninth-grade restructure during the 2019–20 school year due to the implementation of the BARR model?

9. To what extent do you feel the BARR ninth-grade structure changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

D. Team Meetings (5 minutes)

10. In what ways, if any, did you use student-level data (alone or with other teachers) to guide your instructional decisions during the 2019–20 school year?
    a. How important was it for you to be able to access real-time student-level data during the day?

11. Were you part of a BARR team structure at your school during the 2019–20 school year? Did you share the same students across core classes?
    a. Did the BARR team have regular team meetings? If so, how often?
        i. Was this the same before and after school campus closure?
    b. In what ways, if any, were you able to discuss student progress on your BARR team?

12. To what extent do you feel having BARR team meetings changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

13. Were there any particular aspects of the BARR team meetings that you thought helped you as a teacher during the 2019–20 school year? Please provide examples.
    a. Collaborating with other teachers? (Probe: Instructional decisions)
    b. Building positive peer-to-peer relationships? (Probe: Feel respect, trust)

14. Were there any aspects of BARR team meetings you found challenging or difficult during the 2019–20 school year? Please provide examples.

E. Whole Student Approach (5 minutes)

15. In what ways, if any, do you think your students expressed their academic and nonacademic needs during the 2019–20 school year?
    a. Probe: Did this differ before and after school campus closure?
    b. Was this the same or different when thinking about your students' academic and nonacademic strengths during the 2019–20 school year?
        i. Probe: Did this differ before and after school campus closure?

16. Please describe some of the practices you used during the 2019–20 school year:
    a. To address students' academic needs or nonacademic needs.
    b. To build on students' academic strengths or nonacademic strengths.

17. To what extent do you feel the way you addressed students' needs and strengths changed during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.
    b. If so, to what do you attribute this change in the 2019–20 school year?

F. Social-Emotional Learning/Youth Development (5 minutes)

18. Did you facilitate any I-Times in your classroom during the 2019–20 school year? Were there any other social-emotional learning activities or lessons you used?
    a. Were there any particular I-Time activities or lessons that were helpful to you? Please provide examples.
    b. Were you able to conduct I-Times during the school campus closure?

19. To what extent do you feel facilitating I-Times during the 2019–20 school year changed your teaching practice compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

20. In what ways, if any, do you think incorporating I-Times affected:
    a. Developing positive teacher-to-student relationships?
    b. Encouraging positive student-to-student relationships?

G. Student Referral Process (5 minutes)

21. Please describe your experience with the process for referring students for additional supports and services (e.g., risk review, Student Assistance Team) during the 2019–20 school year.
    a. What was the level of feedback or follow-up you received from this process?
    b. [If risk review is not mentioned previously] Did you refer any students to risk review, and what was that experience like?
    c. To what extent did you need to and were you able to refer students to supports and services during the school campus closure?

22. To what extent do you feel the student referral process (i.e., risk review) changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

H. Parent Involvement (5 minutes)

23. Please describe some of the practices or activities you used during the 2019–20 school year to try and foster positive parent-teacher relationships.
    a. Probe: Individually?
    b. As a BARR team?
    c. To what extent, if at all, were you able to foster positive parent-teacher relationships during the school campus closure?

24. To what extent do you feel your practices or activities related to parent involvement changed during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

25. To what extent did you see actual parent involvement change due to the BARR program during the 2019–20 school year compared to the 2018–19 school year? (Better, the same, worse)

I. Contextual Support (5 minutes)

26. What types of BARR support, if any, did the school administration (e.g., principal, assistant principals) provide that helped you during the 2019–20 school year (e.g., attending meetings, supplies, coaching, resources)? Please provide examples.

27. To what extent do you feel the BARR support from school administration (e.g., principal, assistant principals) changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

28. What kinds of BARR support, if any, were provided by other school staff (e.g., counselor, academic coach, department head, BARR coordinator, colleagues) that helped you during the 2019–20 school year? Please provide examples.
    a. Do you think the BARR coordinator's role was an important one? Why or why not?

29. To what extent do you feel that having BARR support from other school staff (e.g., counselor, academic coach, department head, BARR coordinator, colleagues) changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 year.

30. Please describe any support provided to you directly by the BARR program staff (e.g., BARR coach) during the 2019–20 school campus closure.

J. Overall Experience (5 minutes)

31. Compared to ninth-grade students you have had in the past, what do you think the school experience was like for the BARR students you had during the 2019–20 school year?
    a. Probe: What do you think the levels of achievement were like for the BARR students you had during the 2019–20 school year (better, the same, worse)? Attendance (better, the same, worse)? Behavior (better, the same, worse)?
    b. What do you think the school experience after the school campus closure was like for the BARR students you had during the 2019–20 school year?

32. Thinking about any other major external (e.g., neighborhood, community) or internal challenges that impacted your school during the 2019–20 school year, other than COVID-19, did any of them pose major challenges or barriers specific to implementing BARR at your school? (Probe for any aspects of the BARR model that may not have been discussed yet.)

33. Thinking about your overall experience during the 2019–20 school year with BARR, both before and after the school campus closure, is there something that stands out above everything else? Please explain.

34. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

B2.6. Cohort 3 Control Teacher Interview

Hello, my name is _____ and I work for the American Institutes for Research (AIR). I am speaking with you today as part of our work evaluating the nationwide implementation of a ninth-grade program, and your school serves as one of our control schools.

In general, we'll ask about your experiences for the entire 2019–20 school year. Please note, this is not an evaluation of you personally. We would simply like to hear about your experiences and your opinions related to different aspects of teaching at your school during the 2019–20 school year.

This discussion should take no more than the hour we have allotted. Please take a moment to review the interview consent form if you have not done so already. There is more detailed information about this study and your rights as a participant. Please feel free to stop me at any time if you'd like to take a break, reschedule the conversation, or prefer to discontinue the conversation. You can always choose not to answer specific questions. Do you have any questions before we get started?

I would like to be able to audio record our interview in order to accurately capture everything you tell me. The recording is purely for AIR's purposes and will not be shared with anyone else. Do I have your permission to record this interview*?* [If **yes**, begin recording. If **no**, take detailed notes of responses.]

Note: In the questions below, we would like to get teachers' perspectives pre-COVID and post-COVID. Please probe for the other if the teacher does not specifically mention both.

A. Opening Questions (2 minutes)

1. As background, what subject(s) and grade level(s) did you teach during the 2019–20 school year?

2. How long had you been a teacher at this school during the 2019–20 school year?
   a. How long had you been a teacher in total?

3. Could you please describe your personal experience teaching ninth graders during the 2019–20 school year, prior to school closure due to COVID-19?
   a. After the school campus closure?

B. Professional Development (5 minutes)

4. What professional development opportunities did you participate in during the 2019–20 school year prior to the school campus closure?
   a. During the school campus closure, did you participate in any trainings related to school closures/virtual learning?

*For the next question and subsequent questions throughout the interview I will ask you to quantify your response by saying major change, moderate change, minor change, or no change.

5. To what extent do you feel the professional development provided changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)

a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

6. Were there any aspects of the professional development that were particularly helpful to you as a teacher during the 2019–20 school year? Please provide examples (i.e., delivery method of professional development).

7. Were there any aspects of the professional development that you found confusing or difficult to integrate into your teaching, or understand?

C. Ninth-Grade Structure (5 minutes)

8. Please describe the ninth-grade structure at your school prior to the school campus closure during the 2019–20 school year. (Prompts: Related to teacher teams, common planning times, class size, student leveling)
   a. Probe: Did any changes occur during the 2019–20 school year to the ninth grade?

9. To what extent do you feel the ninth-grade structure changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
   a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

D. Team Meetings (5 minutes)

10. In what ways, if any, did you use student-level data (alone or with other teachers) to guide your instructional decisions during the 2019–20 school year?
    a. How important was it for you to be able to access real-time student-level data during the day?

11. Were you part of a team structure at your school during the 2019–20 school year? Did you share the same students across core classes? Did you have departmental teams? (If on multiple teams, probe for each.)
    a. Did you have regular team meetings? If so how often?
        i. Was this the same before and after school campus closure?
    b. In what ways, if any, were you able to discuss student progress on your team?

12. To what extent do you feel having team meetings changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

13. Were there any particular aspects of team meetings that you thought helped you as a teacher during the 2019–20 school year? Please provide examples.
    a. Collaborating with other teachers? (Probe: Instructional decisions)
    b. Building positive peer-to-peer relationships? (Probe: Feel respect, trust)

14. Were there any aspects of team meetings that you found challenging or difficult during the 2019–20 school year? Please provide examples.

E. Whole Student Approach (5 minutes)

15. In what ways, if any, do you think your students expressed their academic and nonacademic needs during the 2019–20 school year?
    a. Probe: Did this differ before and after school campus closure?
    b. Was this the same or different when thinking about your students' academic and nonacademic strengths during the 2019–20 school year?
        i. Did this differ before and after school campus closure?

16. Please describe some of the practices you used during the 2019–20 school year:
    a. To address students' academic needs or nonacademic needs.
    b. To build on students' academic strengths or nonacademic strengths.

17. To what extent do you feel the way you addressed students' needs and strengths changed during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.
    b. If so, to what do you attribute this change in the 2019–20 school year?

F. Social-Emotional Learning/Youth Development (5 minutes)

18. Did you facilitate any social-emotional learning activities or lessons in your classroom during the 2019–20 school year?
    a. Were there any particular activities or lessons that were helpful to you? Please provide examples.
    b. Were you able to conduct social-emotional learning activities or lessons during the school campus closure?

19. To what extent do you feel facilitating social-emotional learning activities or lessons during the 2019–20 school year changed your teaching practice compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

20. In what ways, if any, do you think incorporating social-emotional learning activities or lessons affected:
    a. Developing positive teacher-to-student relationships?
    b. Encouraging positive student-to-student relationships?

G. Student Referral Process (5 minutes)

21. Please describe your experience with the process for referring students for additional supports and services (e.g., Student Assistance Team) during the 2019–20 school year.
    a. What was the level of feedback or follow-up you received from this process?

b. To what extent did you need to and were you able to refer students to supports and services during the school campus closure?

22. To what extent do you feel the student referral process changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

H. Parent Involvement (5 minutes)

23. Please describe some of the practices or activities you used during the 2019–20 school year to try and foster positive parent-teacher relationships.
    a. Probe: Individually?
    b. As a team?
    c. To what extent, if at all, were you able to foster positive parent-teacher relationships during the school campus closure?

24. To what extent do you feel your practices or activities related to parent involvement changed during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

25. To what extent did you see actual parent involvement change during the 2019–20 school year compared to the 2018–19 school year? (Better, the same, worse)

I. Contextual Support (5 minutes)

26. What types of support, if any, did the school administration (e.g., principal, assistant principals) provide that helped you during the 2019–20 school year (e.g., attending meetings, supplies, coaching, resources)? Please provide examples.

27. To what extent do you feel the support from school administration (e.g., principal, assistant principals) changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

28. What kinds of support, if any, were provided by other school staff (e.g., counselor, academic coach, department head, colleagues) that helped you during the 2019–20 school year? Please provide examples.

29. To what extent do you feel having support from other school staff (e.g., counselor, academic coach, department head, colleagues) changed your teaching practice during the 2019–20 school year compared to the 2018–19 school year? (Major Change, Moderate Change, Minor Change, No Change)
    a. Probe: If so, please provide an example of what changed for you during the 2019–20 school year.

30. Please describe any support provided to you directly by district or school staff during the 2019–20 school campus closure.

J. Overall Experience (5 minutes)

31. Compared to ninth-grade students you have had in the past, what do you think the school experience was like for the ninth-grade students you had during the 2019–20 school year?
    a. Probe: What do you think the levels of achievement were like for the ninth-grade students you had during the 2019–20 school year (better, the same, worse)? Attendance (better, the same, worse)? Behavior (better, the same, worse)?
    b. What do you think the school experience after the school campus closure was like for the ninth-grade students you had during the 2019–20 school year?

32. Were there any other major external (e.g., neighborhood, community) or internal challenges that impacted your school during the 2019–20 school year, other than COVID-19?

33. As a control school in this study, your school is likely implementing the BARR (Building Assets, Reducing Risks) model in ninth grade during the 2020–21 school year. Have you heard any information about this program?
    a. In the 2019–20 school year, if given the choice, would you have chosen to be a BARR teacher this year? Why or why not?

34. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

B3. Administrator Interview Protocols

B3.1. Cohort 1 BARR and Control Administrator Interviews

Hello, I am (insert name). As you may recall, our company, the American Institutes for Research (AIR), is leading the scale-up study of the i3 BARR program, which seeks to understand the impact of the BARR model on students and teachers in schools across the country.

Your participation is voluntary, you can stop at any time, and your information will be kept confidential. We want to assure you these conversations will be used solely for research purposes and that nothing you say here will be attributed to you or reported to others within your school.

We would like to record this discussion to make sure we do not miss anything in our conversation. This recording will only be used to verify the notes taken during the discussion and will not be shared with anyone outside of the research team. Is it okay if we record this interview?

The interview should take about 30 minutes and consists of open-ended questions that will allow you to tell your story. (Share consent form with participant.) Do you agree to participate in this study? Thank you. Do you have any questions before we begin?

Treatment	Control
The BARR model was implemented in your school this year. These questions will focus on your school's experience with BARR implementation. **BARR Implementation** 1. Can you give us an overview of what changes have occurred at your school this year since implementing BARR (i.e., master schedules, professional development training, staff - student interactions)? 2. Over the past year, has your school put in place any other new programs we should know about? *a.* Has your school changed or altered any programs in place? **BARR Support and Involvement** 3. Did you or another administrator attend the initial summer/early fall training on the BARR model? *a.* Did you or another administrator attend any other supplemental BARR trainings during the school year?	Your school will be implementing the BARR model next year. These questions will focus on your school's current context and any preparation for BARR activities. **Program Activities** 1. Could you please let us know the progress of any major programs or initiatives that your school implemented this year (e.g., district policies, interventions, supports, structures)? a. Were any of these major programs or initiatives specific to the ninth grade? 2. Over the past year, has your school implemented any new programs since the fall we should know about? 3. Over the past year, has your school changed or altered any programs in place? **BARR Preparation** 4. Has your school set any plans or strategies in place to prepare for BARR implementation next year? a. If yes, please explain. b. If no, do you have a timeline in place for preparation?

Treatment	Control
4. How often did you or another administrator attend the BARR block meetings this year? (Probe for level of participation in meetings.) 5. How often did you or another administrator attend the BARR risk review meetings? (Probe for level of participation in meetings.) 6. How frequently did you or another administrator meet with the BARR coordinator? (Probe for what topics were discussed with the BARR coordinator.) **Reflection on BARR** 7. Since BARR implementation, what has been the reaction of the staff regarding the BARR process? (Probe for the BARR training process and the BARR implementation process.) 8. How would you describe the level of parent involvement related to the BARR team activities (low, moderate, high)? *a.* How does that compare to last year (or to students not on a BARR team)? **Perceived Student Impact** 9. What do you think the school experience has been like for the BARR group this year? *a.* Probe: What do you think the levels of achievement, attendance, and behavior have been like for the group of students in BARR this year (better, the same, worse)? 10. How would you describe your school climate after BARR implementation this year? *a.* Related to learning? *b.* Related to providing student supports (academic, social, emotional)? **BARR Sustainability** Looking ahead, there are 2 more years of BARR implementation planned. 11. Do you have all of the resources (and district support) you need to continue to implement BARR? Why or why not?	5. Do you think you will have all of the resources (and district support) you need to implement the BARR model next year? Why or why not? 6. Have you communicated with staff yet about implementing the BARR model next year? a. What has been the reaction of the staff to the changes related to BARR? 7. What are some of your goals regarding BARR Implementation? a. Do you foresee a change in teacher-student interaction? b. Do you foresee a change in your school climate? 8. Do you foresee a change in the school experience of ninth-grade students compared to this year after BARR implementation? (Probe for achievement, attendance, and behavior.)

Treatment	Control
12. How do you foresee your school sustaining the BARR program after the grant? *a.* What strategies have you thought of or are currently in place to sustain the model? 13. Thinking about your overall experience, is there one thing that stands out above everything else that you would want other principals to know about BARR? Please explain.	
In Closing: Anything Else? 14. Were there any major external or internal challenges that impacted your school this year that you haven't mentioned? *a.* Were there any major challenges or barriers specific to implementing BARR at your school this year? 15. Is there anything else you would like to share with us about the BARR model or the school's future plans?	**In Closing: Anything Else?** 9. Were there any major external or internal challenges that impacted your school this year that you haven't mentioned? 10. Is there anything else you would like to share with us about the BARR model or the school's future plans?

B3.2. Cohort 2 BARR and Control Administrator Interviews

Hello, I am (insert name). As you may recall, our company, the American Institutes for Research (AIR), is leading the scale-up study of the i3 BARR program, which seeks to understand the impact of the BARR model on students and teachers in schools across the country.

Your participation is voluntary, you can stop at any time, and your information will be kept confidential. We want to assure you these conversations will be used solely for research purposes and that nothing you say here will be attributed to you or reported to others within your school.

We would like to record this discussion to make sure we do not miss anything in our conversation. This recording will only be used to verify the notes taken during the discussion and will not be shared with anyone outside of the research team. Is it okay if we record this interview?

The interview should take about 30 minutes and consists of open-ended questions that will allow you to tell your story. (Share consent form with participant.) Do you agree to participate in this study? Thank you. Do you have any questions before we begin?

Treatment	Control
The BARR model was implemented in your school this year. These questions will focus on your school's experience with BARR implementation. **BARR Implementation** 1. Can you give us an overview of what changes have occurred at your school this year since implementing BARR (i.e. master schedules, professional development training, staff-student interactions)? 2. Over the past year, has your school put in place any other new programs we should know about? a. Has your school changed or altered any programs already in place? **BARR Support and Involvement** 3. Did you or another administrator attend the initial summer/early fall training on the BARR model? a. Did you or another administrator attend any other supplemental BARR trainings during the school year? b. Did you or another administrator attend any BARR conferences during the school year?	Your school will be implementing the BARR model next year. These questions will focus on your school's current context and any preparation for BARR activities. **Program Activities** 1. Could you please let us know the progress of any major programs or initiatives that your school implemented this year (e.g., district policies, interventions, supports, structures)? a. Were any of these major programs or initiatives specific to the ninth grade? 2. Over the past year, has your school implemented any new programs since the fall we should know about? 3. Over the past year, has your school changed or altered any programs already in place? **BARR Preparation** 4. Has your school set any plans or strategies in place to prepare for BARR implementation next year? a. If yes, please explain. b. If no, do you have a timeline in place for preparation?

Treatment	Control
4. How often did you or another administrator attend the BARR block/team meetings this year? (Probe for level of participation in meetings.) 5. How often did you or another administrator attend the BARR risk review meetings? (Probe for level of participation in meetings.) 6. How frequently did you or another administrator meet with the BARR coordinator? (Probe for what topics were discussed with the BARR coordinator.) **Reflection on BARR** 7. Since BARR implementation, what has been the reaction of the staff regarding the BARR process? (Probe for the BARR training process and the BARR implementation process.) 8. How would you describe the level of parent involvement related to the BARR team activities (low, moderate, high)? a. How does that compare to last year (or to students not on a BARR team)? **Perceived Student Impact** 9. What do you think the school experience has been like for the BARR group this year? a. Probe: What do you think the levels of achievement, attendance, and behavior have been like for the group of students in BARR this year (better, the same, worse)? 10. How would you describe your school climate after BARR implementation this year? a. Related to learning? b. Related to providing student supports (academic, social, emotional)? **BARR Sustainability** Looking ahead, there are 2 more years of BARR implementation planned. 11. Do you have all of the resources (and district support) you need to continue to implement BARR? Why or why not?	5. Do you think you will have all of the resources (and district support) you need to implement the BARR model next year? Why or why not? 6. Have you communicated with staff yet about implementing the BARR model next year? a. What has been the reaction of the staff to the changes related to BARR? 7. What are some of your goals regarding BARR implementation? a. Do you foresee a change in teacher-student interaction? b. Do you foresee a change in your school climate? 8. Do you foresee a change in the school experience of ninth-grade students compared to this year after BARR implementation? (Probe for achievement, attendance, and behavior.)

Treatment	Control
12. How do you foresee your school sustaining the BARR program after the grant? a. What strategies have you thought of or are currently in place to sustain the model? 13. Thinking about your overall experience, is there one thing that stands out above everything else that you would want other principals to know about BARR? Please explain.	
In Closing: Anything Else? 14. Were there any major external or internal challenges that impacted your school this year that you haven't mentioned? a. Were there any major challenges or barriers specific to implementing BARR at your school this year? 15. Is there anything else you would like to share with us about the BARR model or the school's future plans?	**In Closing: Anything Else?** 9. Were there any major external or internal challenges that impacted your school this year that you haven't mentioned? 10. Is there anything else you would like to share with us about the BARR model or the school's future plans?

B3.3. Cohort 3 BARR Administrator Interview

Hello, I am (insert name). Our company, the American Institutes for Research (AIR), is leading the scale-up study of the i3 BARR program, which seeks to understand the impact of the BARR model on students and teachers in schools across the country. A component of the study is centered on understanding the context of your school and your leadership role to better understand the factors that could affect implementation. As such, the purpose of this interview is to dig deeper into the following school components: context, resource utilization, and BARR implementation for the 2019–20 school year. Our initial interview was planned for last spring, but given the many challenges related to COVID-19, we postponed our data collection. We are still interested in understanding the context for your school and leadership last year. For this reason, our conversation will focus primarily on your reflection of the 2019–20 school year.

Your choice to participate in this interview is completely voluntary, and you may choose to stop at any time. We want to assure you these conversations will be kept confidential and that nothing you say here will be attributed to you or reported to others within your school. We would like to record this interview to make sure we do not miss anything in our conversation. This recording will only be used to verify the notes taken during the interview and will not be shared with anyone outside of the research team. Is it okay if we record this interview?

The interview should take the entire hour we have allotted and consists of open-ended questions that will allow you to tell your story. (Share consent form with participant.) Do you agree to participate in this study? Thank you. Do you have any questions before we begin?

A. Context (10 minutes)

1. As background, how long had you been an administrator at this school during the 2019–20 school year?
   a. How long had you been an administrator in total?

2. What about the history and context of your school do you think made it an ideal candidate for implementing the BARR model during the 2019–20 school year? (Probe: Student population, teachers and staff, surrounding community)

3. Other than BARR, could you please let us know about any major programs or initiatives that your school implemented during the 2019–20 school year (e.g., district policies, interventions, supports, structures)?
   a. Are any of the programs specific to the ninth grade?

B. Resource Utilization (5 minutes)

4. Did you have all of the resources you needed to implement BARR during the 2019–20 school year? Why or why not (e.g., personnel, materials, time, funding)?
   a. Probe: Was this the same before and after school campus closure (e.g., internet/equipment, video conferencing)?

5. Could you please describe the student data management systems you used during the 2019–20 school year?

C. BARR Implementation (5 minutes)

6. Can you give us an overview of what changes occurred at your school during the 2019–20 school year since implementing BARR (i.e., master schedules, professional development training, staff-student interactions)?

D. BARR Support and Involvement (10 minutes)

7. Did you or another administrator attend the **initial foundational** training on the BARR model?
   a. Did you or another administrator attend any other supplemental BARR trainings during the 2019–20 school year?
   b. Did you or another administrator attend any professional development/webinars provided by BARR related to virtual learning?
   c. Overall, how would you describe your experiences with these trainings (e.g., anything especially helpful, not helpful)?

8. How often did you or another administrator attend the **BARR block/team meetings** during the 2019–20 school year?
   a. Did you or another administrator attend **BARR block/team meetings** after school campus closure?
   b. What was your role in these meetings?

9. How often did you or another administrator attend the **BARR risk review meetings** during the 2019–20 school year?
   a. Did you or another administrator attend **BARR risk review meetings** after school campus closure?
   b. What was your role in these meetings?

10. How frequently did you or another administrator meet with the BARR coordinator during the 2019–20 school year?
    a. Did you or another administrator meet with the BARR coordinator after school campus closure?
    b. What were the topics covered?

E. Reflection on BARR (5 minutes)

11. What was the reaction of staff regarding BARR implementation before or during the 2019–20 school year? (Probe for the BARR training process and the BARR implementation process.)
    a. In what ways, if any, do you feel BARR enabled teachers/staff to support students during school campus closure?

12. How would you describe the level of parent involvement related to the BARR team activities during the 2019–20 school year (low, moderate, high)?
    a. How does that compare to the 2018–19 school year (or to students not on a BARR team)?
    b. How would you describe the level of parent involvement after school campus closure?

F. Perceived Student Impact (10 minutes)

13. What do you think were the impacts of the BARR program on the school experience for BARR students in the 2019–20 school year?

a. What were the levels of achievement like for BARR students during the 2019–20 school year compared to prior years (better, the same, worse)? Attendance (better, the same, worse)? Behavior (better, the same, worse)?

G. BARR Sustainability (10 minutes)

Looking ahead, there are 2 more years of BARR implementation planned.

14. Do you have all of the resources (and district support) you need to continue to implement BARR? Why or why not?

15. How do you foresee your school sustaining the BARR program after the 3-year grant period?
    a. What strategies have you thought of or are currently in place to sustain the model?

16. Thinking about your overall experience, is there one thing that stands out above everything else that you would want other principals to know about BARR? Please explain.

H. In Closing: Anything Else? (5 minutes)

17. Did your school reach out to or receive support from BARR in response to the school campus closure due to COVID-19 in the 2019–20 school year? If yes, please describe.

18. Were there any other major external or internal challenges that impacted your school during the 2019–20 school year that you haven't mentioned?
    a. Were there any other major challenges or barriers specific to implementing BARR at your school during the 2019–20 school year?

19. Is there anything else you would like to share with us about the BARR model or the school's future plans?

B3.4. Cohort 3 Control Administrator Interview

Hello, I am (insert name). Our company, the American Institutes for Research (AIR), is leading the scale-up study of the i3 BARR program, which seeks to understand the impact of the BARR model on students and teachers in schools across the country. A component of the study is centered on understanding the context of your school and your leadership role to better understand the factors that could affect implementation. As such, the purpose of this interview is to dig deeper into the following school components: context, resource utilization, and BARR preparation for the 2019–20 school year. Our initial interview was planned for last spring, but given the many challenges related to COVID-19, we postponed our data collection. We are still interested in understanding the context for your school and leadership last year. For this reason, our conversation will focus primarily on your reflections on the 2019–20 school year.

Your choice to participate in this interview is completely voluntary, and you may choose to stop at any time. We want to assure you these conversations will be kept confidential and that nothing you say here will be attributed to you or reported to others within your school. We would like to record this interview to make sure we do not miss anything in our conversation. This recording will only be used to verify the notes taken during the interview and will not be shared with anyone outside of the research team. Is it okay if we record this interview?

The interview should take the entire hour we have allotted and will consist of open-ended questions that will allow you to tell your story. (Share consent form with participant.) Do you agree to participate in this study? Thank you. Do you have any questions before we begin?

A. Context (10 minutes)

1. As background, how long had you been an administrator at this school during the 2019–20 school year?
   a. How long had you been an administrator in total?

2. What about the history and context of your school do you think makes it an ideal candidate for implementing the BARR model in the 2020–21 school year? (Probe: Student population, teachers and staff, surrounding community)

3. Could you please let us know about any major programs or initiatives that your school implemented during the 2019–20 school year (e.g., district policies, interventions, supports, structures)?
   a. Were any of the programs specific to the ninth grade?

B. Resource Utilization (10 minutes)

4. Did you have all of the resources you needed during the 2019–20 school year? Why or why not (e.g., personnel, materials, time, funding)?
   a. Probe: Was this the same before and after school campus closure (e.g., internet/equipment, video conferencing)?

5. Could you please describe the student data management systems you used during the 2019–20 school year?

C. BARR Preparation (15 minutes)

Your school is currently implementing the BARR model this year. These next questions will focus on your school's preparation for BARR activities during the 2019–20 school year.

6. During the 2019–20 school year, did your school have plans in place to prepare for BARR implementation in the 2020–21 school year?
   a. If yes, please explain.

7. Thinking back to the 2019–20 school year, did you think you would have all of the resources (and district support) you needed to implement the BARR model in the 2020–21 school year? Why or why not?

8. In the 2019–20 school year, did you communicate with staff about implementing the BARR model in the 2020–21 school year?
   a. What was the reaction of staff to the changes related to BARR?

9. Did you or another administrator attend any trainings on the BARR model in the 2019–20 school year?
   a. Did you or another administrator attend any professional development/webinars provided by BARR related to virtual learning?
   b. If so, how would you describe your experiences with these trainings (e.g., anything especially helpful, not helpful)?

10. What are some of your goals regarding BARR implementation?
    a. Do you foresee a change in teacher-student interaction?
    b. Do you foresee a change in your school climate?

11. Do you foresee a change in the school experience of ninth-grade students compared to the 2019–20 school year after BARR implementation? (Probe for achievement, attendance, and behavior.)

D. In Closing: Anything Else? (10 minutes)

12. Did your school reach out to or receive support from BARR in response to the school campus closure due to COVID-19 in the 2019–20 school year? If yes, please describe.

    Were there any other major external or internal challenges that impacted your school during the 2019–20 school year that you haven't mentioned?

13. Is there anything else you would like to share with us about the BARR model or the school's future plans?

B4. Teacher Survey Scales and Items

Exhibit B4.1 details the items, stems, and response options for the nine constructs measured on the teacher survey. Exhibit B4.2 includes the COVID-19-related items, stems, and response options added to the Cohort 3 teacher survey. Items marked with a single asterisk (*) were reverse-coded in the analysis. Items marked with two asterisks (**) were excluded from the Cohort 3 teacher survey.

Exhibit B4.1. Teacher Survey Items Administered to Cohorts 1, 2, and 3

Strength-Based Mindset	
Stem	To what extent do you agree or disagree with the following statements?
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. In addition to knowing how my students do academically, I know about their strengths and struggles outside my classroom, including at home and in their community. 2. I know enough about each of my students to be able to support and guide them by leveraging their individual strengths and interests. 3. It is (not) important to my students that I know their personal strengths and interests.[55]* 4. When I acknowledge my students' strengths and interests, it has a positive effect on their performance in other areas. 5. When I advise students, I try to push them to get out of their comfort zone.
Student Academic Motivation	
Stem	How many of the students in your Grade 9 classes:
Response Options	None (1), some (2), about half (3), most (4), nearly all (5)
Items	1. Come to class on time? 2. Attend class regularly? 3. Come to class prepared with the appropriate supplies and books? 4. Regularly pay attention in class? 5. Actively participate in class activities? 6. Always turn in their homework?

[55] In our review of the reliability after Cohort 1, "strength-based mindset" had low internal consistency (Chronbach's alpha = .57). This is primarily due to a poor fitting item. The tone of the item "It is not important to my students that I know their personal strengths and interests" is more negative than the tone of other items in the scale. When we remove this item from the scale, the Rasch reliability and Cronbach's alpha improve (.63 and .69, respectively) for Cohort 1. For Cohorts 2 and 3, we improved this scale by revising the poor fitting item to read "It is important to my students that I know their personal strengths and interests."

Perception of Students' Behavior, Commitment, and Attitudes	
Stem	Most students in your Grade 9 classes:
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. Push themselves to meet high standards. 2. Put off doing things they don't like to do.* 3. Do just enough to get by on their schoolwork.* 4. Set goals for doing better in school and keep track of whether they are improving. 5. Take pride in the quality of their work. 6. Work hard to overcome their challenges. 7. Take initiative to get things done without being asked or reminded. 8. Are open to suggestions for improvement.

View of Students' Observed Behavior, Commitment, and Attitudes	
Stem	Most students in your Grade 9 classes:
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. Are willing to help when they see someone having a problem. 2. Work well with students from different ethnic, religious, cultural, or political backgrounds. 3. Try to stop their friends from spreading rumors or gossiping about others. 4. Treat teachers and staff with respect, even when they disagree with them. 5. Take an active role in helping solve school problems. 6. Encourage each other to follow the rules. 7. Help others on schoolwork without letting them copy or cheat.

Use of Data	
Stem	To what extent do you agree or disagree with the following statements for your Grade 9 classes?
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. I feel comfortable using data from my own Grade 9 classes to monitor and analyze my students' performance. 2. I have access to data from other Grade 9 classes to help me make decisions about what students need and how I can help them. 3. I have the tools I need to target specific interventions to my students when they need them.

	4. I do not have enough time to use the data I have effectively.* 5. I often talk with other teachers about performance and attendance data for my Grade 9 students. 6. It is critical to have a complete picture of my students' performance to do my job properly. 7. I have adequate support for the effective use of available data on my students.
Teacher Self-Efficacy	
Stem	How much can you do to…
Response Options	Nothing or very little (1), some (2), a fair amount (3), a great deal (4)
Items	1. Overcome the influence of adverse community conditions on students' learning? 2. Promote learning when there is lack of support from the home? 3. Control disruptive behavior in the classroom?** 4. Motivate students who show low interest in schoolwork? 5. Get through to the most difficult students?[56] 6. Get students to work together? 7. Keep students on task on difficult assignments?** 8. Get students to do their [Cohorts 1 and 2: homework; Cohort 3: assignments]?[57]
Collaboration With and View of Colleagues	
Stem	To what extent do you agree or disagree with the following statements about the other Grade 9 teachers with whom you work?
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. We work together to coordinate our approach to addressing student needs. 2. We meet regularly to discuss the performance of individual students. 3. We meet regularly with counselors and school administrators to discuss the performance and needs of individual students. 4. We trust each other. 5. We really care about each other.

[56] This item was changed to "Get through to the most *disengaged* students" for Cohort 3 when the survey was administered when school campuses were still closed.
[57] This item was changed to "Get students to do their *assignments?"* for Cohort 3 when the survey was administered when school campuses were still closed.

	6. We feel jointly responsible that all students learn. 7. We help maintain discipline in the whole school, not just our own classroom. 8. I wish I had more time to meet with other teachers about my Grade 9 students.*
View of the School's Supports Provided to Teachers	
Stem	To what extent do you agree or disagree with the following statements about your school? My school…
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. Is a supportive and inviting place for students to learn. 2. Provides effective counseling and support services for students. 3. Promotes trust and collegiality among staff. 4. Provides me with the materials, resources, and training I need to do my job effectively.
Postsecondary Educational Expectations	
Stem	To the best of your knowledge, what percentage of students at your school:
Response Options	None (1), some (2), about half (3), most (4), nearly all (5)
Items	1. Will graduate from high school ready for postsecondary education without the need for remedial coursework? 2. Will drop out of postsecondary education before graduating? 3. Will graduate from postsecondary education?
Stem	To what extent do you agree or disagree with the following statement?
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. Teachers in this school feel that it is a part of their job to prepare students to succeed in college.

Exhibit B4.2. Additional COVID-19 School Closure Survey Items Administered to Cohort 3 Teachers

Expectations for Assignment Completion During School Closure	
Stem	**Were students expected to complete class assignments during the campus closure last school year (2019–20)? (Select one.)**
Response Options	**Yes, students were expected to complete assignments (1); No, students were sent assignments but were not expected to complete them (2); No, students were not sent assignments while the school campus was closed (3)**
Technology Use During School Closure	
Stem	To what extent did you use the following technologies to interact or communicate with your Grade 9 students while your campus was closed last school year (2019–20)?
Response Options	Not at all (1), to a small extent (2), to a moderate extent (3), to a great extent (4)
Items	1. Phone or voice calls 2. E-mail 3. Texting or instant messaging 4. Educational software platform (e.g., Google Classroom, Canvas, Blackboard Learn) 5. Social networking (e.g., Facebook, Twitter, ClassDojo) 6. Video conferencing (e.g., Zoom, GoToMeeting) 7. Other (please specify):
Frequency of Communication With Students	
Stem	While your campus was closed last school year (2019–20), how often did you communicate with your Grade 9 students?
Response Options	Never or almost never (1), sometimes (2), most school days (3), almost every school day (4), every school day (5)
Student Internet Needs During School Closure	
Stem	How often did your Grade 9 students need to access the internet to participate in school while your campus was closed last school year (2019–20)?
Response Options	Never or almost never (1), sometimes (2), most school days (3), almost every school day (4), every school day (5)

Student Devices and Services Provided During School Closure	
Stem	What devices and services did the school provide to Grade 9 students to participate in school while your campus was closed last school year (2019–20)? (Select all that apply.)
Response Options	No devices or services were provided (1), desktop computer (a personal computer that is too large or heavy to be frequently moved from place to place) (2), laptop or notebook (3), tablet or e-book reader (4), smartphone (5), mobile data plan (e.g., hotspot, Verizon, T-Mobile) (6), internet services (e.g., Spectrum, Xfinity, AT&T) (7), other (please specify) (8)
Comfort With Technology	
Stem	How comfortable were you with using the following technologies to teach students while your campus was closed last school year (2019–20)?
Response Options	Not comfortable (1), somewhat comfortable (2), moderately comfortable (3), very comfortable (4)
Items	1. Online teaching using video-based platforms (e.g., Zoom, GoToMeeting) 2. Interactive whiteboards 3. Social networking (e.g., Facebook, Twitter, ClassDojo) 4. Class blogs and wikis 5. Podcasts 6. Mobile learning apps 7. Learning management platform (e.g., Canvas) 8. Virtual field trips (e.g., Google Arts & Culture) 9. Other (please specify):
Self-Efficacy for Instructional Use of Technology	
Stem	Indicate how strongly you agree or disagree with the following statements.
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. I learn technology easily. 2. When I am confronted with a technology-related problem, I usually find good solutions. 3. I easily find new technologies to meet my teaching goals. 4. I feel confident managing a classroom where students are using technology.

Communication With Colleagues During School Closure	
Stem	While your campus was closed last school year (2019–20), how often did you communicate with other teachers?
Response Options	Never or almost never (1), sometimes (2), most school days (3), almost every school day (4), every school day (5)
Open-Ended Follow-Up Question	Please describe what you talked to other teachers about while campus was closed last school year (2019–20).
School Supports During School Closure	
Stem	To what extent did you receive the support you needed from your school to successfully teach outside the classroom while campus was closed last school year (2019–20)?
Response Options	Not at all (1), to a small extent (2), to a moderate extent (3), to a great extent (4)
Open-Ended Follow-Up Question	Please provide additional feedback about your experience with online teaching while your campus was closed last school year (2019–20).

B5. Student Survey Scales and Items

Exhibit B5.1 details the items, stems, and response options for each of the six constructs measured on the student survey. Exhibit B5.2 includes the COVID-19-related items, stems, and response options added to the Cohort 3 student survey.

Exhibit B5.1. Student Survey Items Administered to Cohorts 1, 2, and 3

Supportive Relationships	
Stem	**Thinking about your English, science, and social studies teachers, for how many of those teachers is each statement true?**
Response Options	**None (1), one of my teachers (2), two of my teachers (3), three of my teachers (4)**
Items	1. If my teacher asks me how I am doing, I often feel that they are really interested in my answer. 2. My teacher is interested in my career after I finish school. 3. If I came back to visit class 3 years from now, my teacher would be excited to see me. 4. My teacher is interested in what I do outside of class. 5. If I walked into class upset, my teacher would be concerned. 6. If I had something on my mind, my teacher would carefully listen to me.
Expectations and Rigor	
Stem	Still thinking about your English, math, science, and social studies classes, for how many of these classes is each statement true?
Response Options	None of my classes (1), one of my classes (2), two of my classes (3), three of my classes (4)
Items	1. My teacher often encourages me to do my best. 2. My teacher often takes the time to make sure I understand the material. 3. My teacher doesn't let people give up when the work gets hard. 4. My teacher accepts nothing less than my full effort. 5. My teacher pushes me to become a better thinker. 6. My teacher makes us try to find the answers on our own before he or she answers our questions. 7. I learn a lot from feedback on my work. 8. The classwork helps me learn the course materials. 9. The work we do in class is good preparation for tests. 10. I know what my teacher wants me to learn. 11. It's clear what I need to do to get a good grade.

Student Engagement	
Stem	Still thinking about your English, math, science, and social studies classes, for how many of these classes is each statement true?
Response Options	None of my classes (1), one of my classes (2), two of my classes (3), three of my classes (4)
Items	1. I often get so focused on class activities that I lose track of time. 2. I am interested in the class. 3. When I am not in class, I often talk about ideas from class. 4. I often participate in these classes. 5. I am often excited to go to my classes. 6. I always prepare for class. 7. I ask questions when I don't understand the lesson.
School Climate (inclusive of emotional safety)	
Stem	How strongly do you agree or disagree with the following statements about your high school?
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. I feel like I belong. 2. Students at this school get along well with each other. 3. At this school, students work on listening to others to understand what they are trying to say. 4. I am happy to be at this school. 5. I feel like I am part of this school. 6. I feel socially accepted. 7. I feel safe at this school. 8. Adults working at this school treat all students respectfully.
Schoolwide Future Orientation	
Stem	How strongly do you agree or disagree with the following statements about your high school?
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. My teachers make sure that all students are planning for life after graduation. 2. My teachers work hard to make sure that all students are learning. 3. High school is seen as preparation for the future.

	4. All students are encouraged to pursue further education. 5. My teachers pay attention to all students, not just the top students. 6. My teachers work hard to make sure that students stay in school.
Student Self-Efficacy	
Stem	How often are the following statements true about you?
Response Options	Never or almost never true (1), sometimes true (2), usually true (3), always or most always true (4)
Items	1. I believe I will be able to reach my goals. 2. I know I can complete difficult tasks. 3. I believe I can do whatever I decide to do. 4. I believe I will be able to overcome challenges. 5. I know I can do many different things well. 6. Compared to most other people, I can do most tasks very well. 7. Even when things are tough, I can perform quite well.

Exhibit B5.2. Additional COVID-19 School Closure Survey Items Administered to Cohort 3 Students

Expectations for Assignment Completion During School Closure	
Stem	Like many other schools across the nation, your school campus closed during your ninth-grade year due to COVID-19. The following questions are about the technology you had access to while your school campus was closed during your ninth-grade year and your educational experience during that time. Were you expected to complete class assignments while your school campus was closed during your ninth-grade year?
Response Options	Yes, I was expected to complete assignments (1); No, teachers sent me assignments, but I was not expected to complete them (2); No, I was not sent assignments while the school campus was closed (3)
Student Access to Technology During School Closure	
Stem	How often did you have access to the following technologies to participate in school while the school campus was closed during your ninth-grade year?
Response Options	Never or rarely (1), sometimes (2), often (3), always (4)
Items	1. Internet or a data plan 2. Device (e.g., laptop, desktop, tablet, smartphone)

Use of Technology During School Closure	
Stem	To what extent did you use the following technologies to interact or communicate with your teachers while your school campus was closed during your ninth-grade year?
Response Options	Not at all (1), to a small extent (2), to a moderate extent (3), to a great extent (4)
Items	1. Phone or voice calls 2. E-mail 3. Texting or instant messaging 4. Educational software platform (e.g., Google Classroom, Canvas, Blackboard Learn) 5. Social networking (e.g., Facebook, Twitter, ClassDojo) 6. Video conferencing (e.g., Zoom, GoToMeeting) 7. Other (please specify):
Frequency of Communication With Teachers During School Closure	
Stem	In general, how often did you interact or communicate with your teachers while your school campus was closed during your ninth-grade year?
Response Options	Never or almost never (1), sometimes (2), most school days (3), almost every school day (4), every school day (5)
Student Engagement Outside the Classroom	
Stem	How strongly do you agree or disagree with the following statements about continuing school outside of the classroom while your school campus was closed during your ninth-grade year?
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. I wanted to continue learning even though the school campus was closed. 2. I completed assignments my teachers sent me. 3. My family actively helped me continue my schoolwork. 4. I was able to focus on my class assignments. 5. I reached out to my teachers when I didn't understand the lesson or assignment. 6. I talked to my friends about what I was learning even though the school campus was closed.

Supportive Relationships Outside the Classroom	
Stem	How strongly do you agree or disagree with the following statements about your high school while your school campus was closed during your ninth-grade year?
Response Options	Strongly disagree (1), disagree (2), agree (3), strongly agree (4)
Items	1. I felt connected to my school even though the school campus was closed. 2. My teachers stayed in touch with me even though the school campus was closed. 3. My teachers were interested in how I was doing even though the school campus was closed. 4. My teachers cared about how the school campus closure affected me.
Open-Ended Follow-Up Question	Please provide additional feedback about your experience with distance learning while your school was closed during your ninth-grade year (2019–20). What worked well and what did not work well?

Appendix C. Implementation

Fidelity Measurement

We measured implementation fidelity using ratings from interviews with BARR coordinators and site visit observations of BARR activities (block meetings, I-Time lessons, and risk review meetings).[58] Using these interview and observation data, we calculated fidelity scores for each of the eight strategies for each school and then compared the scores with a predetermined threshold for assessing adequacy of model implementation for cohorts 1, 2, and 3.

During the interviews, the BARR coordinators shared and explained their ratings for the school on each indicator for each of the eight BARR components. Whenever possible, two evaluators observed and rated three BARR activities during spring site visits using fidelity rubrics that were first developed for the 2010 i3 development project and subsequently revised during the 2014 i3 validation study.

Ratings for each fidelity indicator from the BARR coordinator structural review interview and the observed BARR activities were assigned using a scale of 1 to 7, corresponding to the following levels of implementation:

Not Yet (1–2): Activity has not been implemented or has been implemented to a limited extent (e.g., preliminary planning or discussions have occurred, but no concrete planning is in place; activities have occurred sporadically, if at all, and at less than one third of the expected frequency for full implementation). In the case of shifts in practice, this is occurring with few, if any, teachers or staff.

Emerging (3–5): Activity has been implemented in an emerging manner so that concrete planning and, in some cases, activities have occurred, but not to the full extent as intended for the BARR model (e.g., planning has begun and concrete details, such as schedules or processes, have been established and finalized; some activities have been implemented, but they represent only a small proportion of the total number of activities that should be in place, such as monthly meetings rather than weekly meetings). In the case of shifts in practice, this may occur with some, but not all, teachers or staff, or it may occur with all teachers or staff but at a superficial or beginning level.

In Place (6–7): Activity has been implemented to a high degree, with all or nearly all anticipated activities completed as intended. In the case of shifts in practice, this has occurred with all or nearly all teachers or staff and has occurred nearly completely to the extent intended.

[58] Because of school closures due to COVID-19, we were unable to conduct site visit observations of BARR activities in Cohort 3. We measured implementation fidelity for Cohort 3 using ratings from interviews conducted with BARR coordinators in fall 2020 and the most recent implementation fidelity scores collected by BARR staff during the 2019–20 coaching site visits.

After the structural component review interview, we combined BARR coordinators' ratings with observation data. For the three observed BARR activities (block meetings, I-Time lessons, and risk review), the final component fidelity rating is based on an equal weight of the rating from the BARR coordinator (50%) and from the average score of observations made during the site visit (50%). All other fidelity ratings are based on interviews with the BARR coordinators.[59]

Setting Thresholds for Measurement

Our assessment of implementation fidelity occurred in the spring of the first year of BARR implementation in each school.[60] These assessments were based on predetermined thresholds for adequacy of model implementation for each of the BARR key components/strategies (e.g., for the professional development indicator, a school would have had to score a 4.0 or higher to be considered as implementing with adequate fidelity). The fidelity thresholds for indicators under each key component were determined based on theoretical expectations and practical experiences at the beginning of the study, in consultation with the program developers. To meet school-level fidelity, schools were expected to meet four of the five priority component thresholds and two of the three additional component thresholds. To meet cohort-level fidelity, two thirds of the schools in a cohort were expected to meet school-level fidelity. Exhibit C1 details the number of indicators and component thresholds. For more information on the indicators under each component, please contact BARR directly.

Exhibit C1. Number of Indicators and Thresholds Used to Calculate Fidelity

BARR Key Component	Indicators	Threshold
Professional Development	3	4.0 or higher
Restructuring Ninth Grade	7	5.3 or higher
Whole Student Emphasis	4	5.3 or higher
Block/Team Meetings	8	5.3 or higher
I-Time	8	5.0 or higher
Risk Review	9	4.7 or higher
Contextual Support	7	4.0 or higher
Family Involvement	5	3.6 or higher

Calculating Fidelity Scores

To calculate the component score for professional development, restructuring, whole student emphasis, contextual support, and family involvement, we averaged the BARR coordinator

[59] For Cohort 3, all fidelity ratings were based on an equal weight of the rating from the BARR coordinator (50%) and the fidelity rating assigned by BARR staff during their final coaching site visit during the 2019–20 school year before school closures (50%).
[60] For Cohort 3, we interviewed BARR coordinators in fall 2020 after the first year of implementation due to school campus closures in spring 2020.

indicator ratings for each component. The BARR coordinator component score for each school was compared against the predetermined threshold for adequate implementation. If the BARR coordinator component score matched or exceeded the threshold, school-level implementation of the component was deemed adequate. If it did not meet the threshold, school-level implementation was deemed inadequate for the component. Exhibit C2. Provides details on the components, indicators and thresholds used to determine implementation fidelity.

Exhibit C2. Fidelity Matrix

Component	Data Source[61]	Indicator-Level Metric	Threshold for Adequate Implementation for a Component	Threshold for Adequate Implementation for a School	Threshold for Adequate Implementation for a Cohort/Program
Priority Components				School meets four of the five priority component thresholds and two of the three additional component thresholds	Two thirds (67%) of schools rated as having adequate school-level implementation on four of the five priority components and two of the three additional components
Restructuring	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if *Not yet present*, 3–5 if *Emerging*, 6–7 if *In Place*	Mean of seven indicators at or above 5.3		
I-Time	BARR coordinator interview, observation data	Up to 7 points total for each indicator; 1–2 if *Not yet present*, 3–5 if *Emerging*, 6–7 if *In Place*	Mean of eight indicators and observations at or above 5.0		
Block Meeting	BARR coordinator interview, observation data	Up to 7 points total for each indicator; 1–2 if *Not yet present*, 3–5 if *Emerging*, 6–7 if *In Place*	Mean of eight indicators and observations at or above 5.3		
Risk Review	BARR coordinator interview,	Up to 7 points total for each indicator; 1–2 if *Not yet present*, 3–5 if	Mean of nine indicators and observations		

[61] Because of school closures due to COVID-19, we were unable to conduct site visit observations of BARR activities in Cohort 3. We measured implementation fidelity for Cohort 3 using ratings from interviews conducted with BARR coordinators in fall 2020 and the most recent implementation fidelity scores collected by BARR staff during the 2019–20 coaching site visits.

Component	Data Source[61]	Indicator-Level Metric	Threshold for Adequate Implementation for a Component	Threshold for Adequate Implementation for a School	Threshold for Adequate Implementation for a Cohort/Program
	observation data	*Emerging*, 6–7 if *In Place*	at or above 4.7		
Professional Development	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if *Not yet present*, 3–5 if *Emerging*, 6–7 if *In Place*	Mean of three indicators at or above 4.0		
Additional Components					
Family Involvement	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if *Not yet present*, 3–5 if *Emerging*, 6–7 if *In Place*	Mean of five indicators at or above 3.6		
Whole Student Emphasis	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if *Not yet present*, 3–5 if *Emerging*, 6–7 if *In Place*	Mean of four indicators at or above 5.3		
Contextual Support	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if *Not yet present*, 3–5 if *Emerging*, 6–7 if *In Place*	Mean of seven indicators at or above 4.0		

For the three observed BARR activities—block meetings, I-Time, and risk review—the final component fidelity rating was based on an equal weight of the BARR coordinator component score (50%) and observations component score (50%). To calculate the school-level observation score, we averaged across all the indicator ratings. If multiple activities of the same type (e.g., block meeting) were observed in a school, an average school-level observation score was calculated from the multiple observation averages. The score was averaged with the BARR

coordinator component score to create an overall component score for block meetings, I-Time lessons, and risk review. If the overall component score matched or exceeded the threshold, school-level implementation of the component was deemed adequate. If it did not meet the threshold, school-level implementation was deemed inadequate. Exhibit C3 – C5. details whether treatment schools met component- and school-level fidelity by cohort.

Exhibit C3. Cohort 1 Implementation Fidelity Findings

Cohort	School	Priority Components					Additional Components			School-Level Fidelity		
		Restructuring	Professional Development	I-Time	Block Meeting	Risk Review	Family Involvement	Whole Student Emphasis	Contextual Support	Priority Met (4 of 5)	Additional Met (2 of 3)	School-Level Fidelity
1	School A1	✓	✓	✓	–	✓	–	✓	✓	✓	✓	✓
1	School A2	✓	✓	✓	✓	✓	–	✓	✓	✓	✓	✓
1	School A3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
1	School A4	–	✓	✓	✓	✓	–	✓	✓	✓	✓	✓
1	School A5	–	–	–	–	✓	–	–	–	–	–	–
1	School A6	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
1	School A7	–	–	–	–	–	–	–	–	–	–	–
1	School A8	✓	✓	✓	–	✓	✓	–	✓	✓	✓	✓
1	School A9	✓	✓	✓	–	–	✓	–	✓	–	✓	–
1	School A10	✓	–	–	–	–	–	✓	–	–	–	–
1	School A11	✓	✓	✓	✓	✓	–	✓	✓	✓	✓	✓
1	School A12	–	–	–	–	–	✓	✓	–	–	✓	–
1	School A13	✓	✓	✓	✓	✓	✓	✓	–	✓	✓	✓
1	School A19	–	–	–	–	–	–	–	–	–	–	–
1	School A21	–	–	–	–	–	–	–	–	–	–	–
1	**Total number of schools with adequate implementation (*n*=15 schools)**	**9**	**9**	**9**	**6**	**9**	**6**	**9**	**8**	**8 (53%)**	**10 (67%)**	**8 (53%)**

Note: A checkmark means that the school met component expectations for adequate implementation. A dash means that the school did not meet component expectations for adequate implementation.

Exhibit C4. Cohort 2 Implementation Fidelity Findings

Cohort	School	Priority Components					Additional Components			School-Level Fidelity		
		Restructuring	Professional Development	I-Time	Block Meeting	Risk Review	Family Involvement	Whole Student Emphasis	Contextual Support	Priority Met (4 of 5)	Additional Met (2 of 3)	School-Level Fidelity
2	School B1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2	School B2	✓	✓	–	–	–	✓	✓	✓	–	✓	–
2	School B3	✓	✓	✓	✓	✓	–	✓	✓	✓	✓	✓
2	School B4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2	School B5	✓	✓	✓	✓	✓	✓	–	✓	✓	✓	✓
2	School B6	–	✓	✓	–	✓	–	–	✓	–	–	–
2	School B7	–	–	–	–	–	–	–	–	–	–	–
2	School B8	–	–	–	–	–	–	–	–	–	–	–
2	School B9	✓	✓	–	–	✓	–	–	✓	–	–	–
2	School B10	✓	✓	–	–	✓	✓	–	✓	–	✓	–
2	School B11	✓	✓	✓	–	✓	✓	✓	✓	✓	✓	✓
2	**Total number of schools with adequate implementation (*n*=11 schools)**	**8**	**9**	**6**	**4**	**8**	**6**	**5**	**9**	**5 (45%)**	**7 (64%)**	**5 (45%)**

Note: A checkmark means that the school met component expectations for adequate implementation. A dash means that the school did not meet component expectations for adequate implementation.

Exhibit C5. Cohort 3 Implementation Fidelity Findings

Cohort	School	Priority Components					Additional Components			School-Level Fidelity		
		Restructuring	Professional Development	I-Time	Block Meeting	Risk Review	Family Involvement	Whole Student Emphasis	Contextual Support	Priority Met (4 of 5)	Additional Met (2 of 3)	School-Level Fidelity
3	School C1	✓	✓	✓	✓	✓	✓	–	✓	✓	✓	✓
3	School C2	–	–	✓	–	✓	–	✓	✓	–	✓	–
3	School C3	–	✓	✓	✓	✓	–	✓	–	✓	–	–
3	School C4	–	✓	–	–	✓	✓	✓	✓	–	✓	–
3	School C5	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
3	School C6	✓	✓	✓	–	✓	✓	–	✓	✓	✓	✓
3	School C7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
3	School C8	✓	✓	✓	–	–	–	✓	✓	–	✓	–
3	School C9	*	*	*	*	*	*	*	*	*	*	*
3	School C10	✓	✓	✓	–	✓	–	✓	✓	✓	✓	✓
3	School C11	–	✓	✓	–	✓	–	–	✓	–	–	–
3	**Total number of schools with adequate implementation (*n*=10 schools)**	**6**	**9**	**9**	**4**	**9**	**5**	**7**	**9**	**6 (60%)**	**8 (80%)**	**5 (50%)**

Note: A checkmark means that the school met component expectations for adequate implementation. A dash means that the school did not meet component expectations for adequate implementation. For Cohort 3, 11 schools were assigned to the treatment condition and were expected to implement the BARR model during the 2019–20 school year. However, we only included 10 schools in our implementation fidelity analysis because we were unable to collect implementation data from one school (noted with *).

Professional Learning Community (PLC) Case Study

The BARR coordinator PLCs were BARR's primary mechanism for reaching schools as school campuses were closing. As a response to the COVID-19 pandemic, BARR scheduled weekly PLC meetings to provide support to schools in their nationwide network. BARR held 10 virtual PLC meetings from March 26, 2020, to May 28, 2020. Prior to March 26, BARR held monthly PLC meetings. School campus closures sidelined in-person work with schools across the country and limited the usual ways of building teacher-to-student and student-to-student relationships and social-emotional learning skills development. BARR sought to mitigate some of these impacts of COVID-19 by leveraging the virtual aspects of the network they had built with schools and identifying ways to emphasize ongoing work focused on building connections and community in the new circumstances.

We worked with BARR to receive permission from spring 2020 PLC attendees to review recordings and documentation related to the PLCs. The purpose of the spring 2020 PLC case study was to capture how BARR and the schools they work with were able to continue to collaborate and address the BARR model goals of building intentional staff-to-staff, staff-to-student, and student-to-student relationships while school campuses were closed in spring 2020.

Methods

We analyzed the PLC Zoom recordings along with attendance files and the chat box log. We reviewed these materials independently and developed summaries of key findings. The team then discussed their individual summaries and developed a master summary. The team asked the following questions for each PLC:

- What was the topic and content delivered? What tools and/or resources were shared?
- What was the format of the session? Was any portion of it peer-led (by a BARR coordinator)? Were there breakout sessions and, if so, for how long?
- Did participants use the chat box to engage with other participants and/or BARR staff? Did the number of attendees (out of those invited) speak to engagement?

Exhibit C6 shows the dates and topics for each of the PLCs.

Exhibit C6. Professional Learning Community Dates and Topics

PLC Session Date	Topic
3/26/2020	Virtual I-Times and Focus on the Whole Teacher: How to Help Your Teachers Feel Connected
4/2/2020	Risk Review and Self-Care
4/9/2020	Conducting a Virtual Block Meeting
4/16/2020	Virtual Work Session With Your BARR Coach and Other BARR Coordinators
4/23/2020	How to Implement the Modified Virtual I-Times and Successful Student Engagement
4/30/2020	Virtual Presenting Workshop
5/7/2020	Family Engagement
5/14/2020	How Do We Finish the Year BARR Strong?
5/21/2020	Guest Speaker: 2010 National Teacher of the Year
5/28/2020	A Celebration of BARR Successes

We also reviewed PLC attendance information to examine the characteristics of the individuals and schools that attended PLC meetings as well as how attendance varied over the course of the spring.[62]

PLC Findings

We found that the PLC work created an opportunity for the BARR staff to (1) create space for adult connection during a turbulent time, and (2) build educator capacity to support students' SEL and stay connected with students and families while school campuses were closed.

Creating Space for Adult Connection. Creating a space for educators across the BARR network to come together to connect and listen to each other was one of the primary foci of the weekly spring 2020 PLC work that we observed. We found that BARR staff opened each PLC by asking participants to use the chat box to share their name, school, and location and respond to an icebreaker question. In this way, BARR staff facilitated connections with and among BARR PLC participants from across the country. BARR staff who attended the PLC meetings also responded to these questions themselves and often read aloud responses to share with the group.

[62] During the April 2 meeting, BARR informed PLC attendees that data were being collected for research purposes and provided an opt-out form for those who did not want to include their data in the study. Because we did not implement this informed consent procedure until the second meeting, the findings presented in this article do not include data from the first PLC on March 26.

Another way in which the BARR staff worked to build a virtual space for participants to connect with each other was by including BARR coordinators (who are school staff) as presenters alongside BARR Center coaches. In four of the PLC meetings that we observed, BARR coordinators presented on BARR components, such as the weekly SEL-focused I-Time lessons and family engagement. Hearing directly from BARR coordinators about their current school implementation experiences created an opportunity for the other BARR coordinators to hear from those in their same position across the country and how they were experiencing and addressing the unprecedented issues that schools faced in spring 2020.

BARR staff employed breakout sessions as another tool to further build a forum for connection and listening. More than half of the PLC meetings we observed included breakout sessions. In general, when included as part of a PLC meeting, breakout sessions comprised about 15% of the total PLC meeting time. Prior to the start of each breakout session, BARR staff posed specific discussion questions to elicit dialogue among BARR PLC participants about the PLC topic and to help participants consider how the topic might relate to their individual school.

Building Educator Capacity to Support Students' SEL and to Stay Connected With Students and Families. Building educator capacity to support students' SEL and to connect with students and families was another area of emphasis that we observed in the spring 2020 weekly PLC work. The BARR team created PLC meetings around topics such as how to adapt the I-Time SEL curriculum to the virtual environment, how to modify student engagement activities for distance learning, and how to connect with students, families, and communities after the sudden closure of school campuses.

The BARR team coordinated peer-led discussions and breakout sessions on virtual I-Time lessons and student engagement strategies informed by the participating BARR coordinators' authentic real-life experiences. The BARR staff reminded participants of the importance of reaching out to students virtually, noting that, "If one person gets something from this, it's critical…as long as you're sending it out, you're offering another point of connection. And they may take it and they may not, and we can't control that…it could be landing in ways that we will never understand." In a separate I-Time discussion, one of the BARR coordinators shared that she used the I-Time lesson as a jumping-off point "to see which students…self-identified as struggling so we could intervene. That was really crucial because that's when we could see the students say, 'I'm super disengaged.' 'Oh, okay, we need to connect you with your counselor who can talk with you for 10 minutes a day.'" The BARR coordinator also shared that students noted during I-Time that "their biggest struggle was organization and motivation," and that then became the focus of subsequent I-Time lessons. BARR PLC participants shared information about helpful practices and asked questions that could inform their future work to support students.

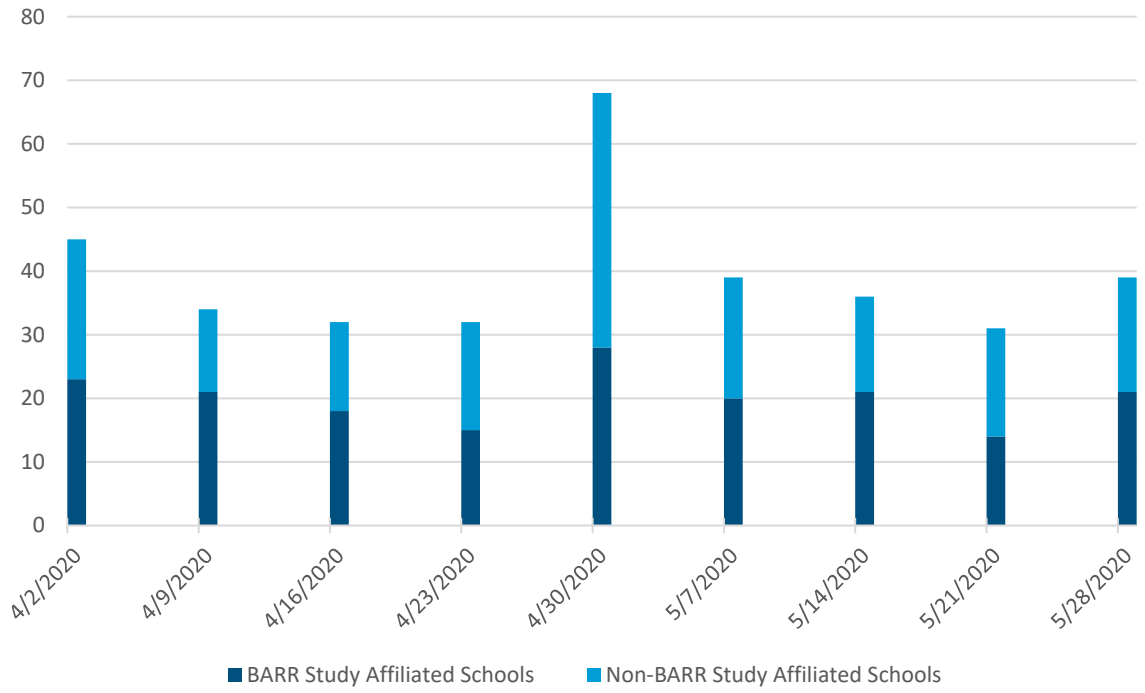
BARR staff also encouraged participants to share what they could do to build connections and relationships with students and their families. One of the BARR coordinators shared that one of the strategies that is working for their ninth-grade students are weekly “town meetings” that have “helped grab kids that we don't normally see,” and staff are able to text students, and say, "Hey, I don't see you on this meeting. Where are you? Come join us." Another BARR coordinator shared that in order to make sure every student is safe, they created weekly, small-group check-ins that initially focused on, "Hi, are you okay?...Do you have internet access?” and evolved into, "Do you need help with math homework? Do you want to play a game with us?” in order to build a point of connection to the school as well as to a caring adult outside the home. BARR coordinators also noted that, with schools closed, they had to increase their outreach to families as well. One BARR coordinator acknowledged, “It's harder for families to drop in and see us, but we still want to be able to engage…A lot of our families don't always come to us, so we come to them…The home visits are continuing.” Another BARR coordinator explained that to reach parents, “we literally sent it [school communication] out on every platform we can think of… We know social media, everyone's looking at it.”

Another way that BARR staff worked to build the capacity of the PLC participants to connect with students, families, and partners was through developing and showcasing tools in collaboration with BARR coordinators, such as a tool to track outreach and contact with students, and a transitions tool. The BARR team shared how the dashboard of the student contact tracker could be used for BARR coordinators and teachers to identify which students had not been in touch with their teachers within the past week. BARR staff also shared a new transitions tool that provides a template for BARR coordinators to use to pull in data about their students and additional information, such as how to best reach specific students and their families, for continued support to students over the summer and into the next grade level.

This case study shows that with an emphasis on relationships, school intervention developers can pivot their work to continue to support schools in a meaningful way during a time of upheaval. As part of the pivot, we interviewed BARR team members about the impetus behind their decision to create weekly PLCs as well as the rationale for the PLC content and structure. BARR staff explained that they intentionally focused on the importance of relationships and trust as well as the tone of communication to ensure that through the PLCs, the schools felt supported and knew that the team was flexible and available to continue to be in a virtual format when the schools were ready. From the BARR team’s perspective, this was particularly important during this period because educators felt as though they had been given a task to do—educate their students outside the classroom setting—without much direction or knowledge about how to proceed.

We found that BARR succeeded in attracting school team members to their PLCs during this time of school campus closures, as nearly 80% of invited individuals attended at least one meeting. In total, 260 individuals from 117 schools were invited to the PLC meetings in April and May 2020, and 206 individuals from 99 schools attended these meetings.[63] Exhibit C7 shows that, overall, the number of attendees at each PLC was consistent throughout the spring, with one exception on April 30 in which attendance was higher than in the other PLCs.[64]

Exhibit C7. Number of Schools Attending by Study and Nonstudy School Status at Each Professional Learning Community



Source: BARR attendance records.

The strong attendance for the PLCs during this time may be due in part to the focus on connections to support students and adults during and beyond the pandemic. By creating space for adult connection and focusing on professional development to build educator capacity to support students' SEL and educator capacity to connect students, families, and partners during this unprecedented time, the BARR team was able to continue to support schools in building intentional staff-to-staff, staff-to-student, and student-to-student relationships.

[63] About 40% of the schools represented at these meetings were previously or currently a part of AIR's evaluation work. These schools had either participated in the i3 BARR Validation Study or are currently participating in the i3 BARR Scale-Up Study.
[64] The PLC held on April 30 was a virtual presentation workshop. It was the highest-attended PLC due in part to the fact that BARR encouraged schools to forward the meeting invitation to any of their contacts with an interest in improving their virtual presentations. This brought additional schools and attendees from outside of BARR's usual network.

Appendix D. Teacher Survey Results by Cohort

Exhibits D1 and D2 present the teacher survey results by cohort.

Exhibit D1. Impacts on Teacher Experiences Measured by Teacher Surveys, by Cohort

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
Cohort 1											
Student academic motivation	146	49.3	9.3	81	47.1	9.5	65	2.12	2.04	0.301	0.22
Teacher collaboration with and view of colleagues	143	55.1	10.7	79	45.0	9.1	64	10.16 ***	2.51	0.000	1.01
Teacher use of data	144	53.6	9.4	80	50.2	11.4	64	3.40 *	1.46	0.021	0.33
Postsecondary educational expectations	146	50.4	8.7	81	45.6	7.7	65	4.80 **	1.41	0.001	0.58
View of the school's supports	143	51.3	8.8	79	50.0	11.3	64	1.29	2.35	0.585	0.13
Teacher self-efficacy	144	53.8	8.3	80	52.8	9.6	64	0.97	1.13	0.393	0.11
Strength-based mindset	147	51.6	8.3	81	49.0	9.6	66	2.53 *	1.24	0.043	0.28
Students' actual behavior	145	48.9	8.7	80	47.0	9.9	65	1.89	1.50	0.211	0.20
Perception of students' behavior	146	48.6	8.9	81	46.3	9.5	65	2.33	1.46	0.114	0.25
Cohort 2											
Student academic motivation	202	49.6	10.8	102	47.2	8.6	100	2.37	2.45	0.337	0.24
Teacher collaboration with and view of colleagues	199	50.6	10.0	99	46.0	9.0	100	4.51 **	1.34	0.001	0.47
Teacher use of data	200	50.0	10.1	100	47.0	9.6	100	3.05 *	1.36	0.026	0.31
Postsecondary educational expectations	205	49.5	11.1	104	48.6	10.6	101	0.91	2.64	0.730	0.08
View of the school's supports	200	49.7	10.5	100	46.1	9.8	100	3.59 *	1.61	0.026	0.35
Teacher self-efficacy	199	51.4	7.8	100	52.2	10.0	99	-0.83	0.84	0.319	-0.09
Strength-based mindset	206	50.9	9.9	105	49.2	9.4	101	1.75	1.04	0.095	0.18
Students' actual behavior	202	49.8	10.3	102	47.3	10.0	100	2.44 **	0.88	0.006	0.24
Perception of students' behavior	202	50.7	9.3	102	46.5	10.2	100	4.19 *	1.64	0.012	0.43

(continues)

Exhibit D1. Impacts on Teacher Experiences Measured by Teacher Surveys, by Cohort (continued)

	Treatment				Control						Effect Size
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	
Cohort 3											
Student academic motivation	158	52.9	9.7	93	54.6	9.9	65	-1.71	1.66	0.307	-0.17
Teacher collaboration with and view of colleagues	153	54.4	9.5	91	47.7	7.9	62	6.75 **	2.41	0.006	0.75
Teacher use of data	156	52.2	9.7	92	46.9	8.0	64	5.30 *	2.04	0.010	0.58
Postsecondary educational expectations	159	52.9	9.1	93	52.7	10.5	66	0.13	1.44	0.927	0.01
View of the school's supports	152	54.1	8.5	90	49.0	9.3	62	5.10 ***	1.39	0.000	0.57
Teacher self-efficacy	153	44.5	10.4	91	44.5	9.6	62	-0.01	0.90	0.988	0.00
Strength-based mindset	159	50.3	11.5	93	48.5	11.0	66	1.73	2.10	0.412	0.15
Students' actual behavior	156	54.6	8.6	92	52.3	10.4	64	2.39	1.68	0.157	0.25
Perception of students' behavior	157	54.2	9.2	93	53.7	10.2	64	0.47	1.36	0.730	0.05
Teacher comfort with technology ‡	158	51.0	10.8	92	48.6	8.9	66	2.39	1.38	0.085	0.24
Self-efficacy for instructional use of technology ‡	158	50.8	10.1	92	48.8	9.9	66	2.00	1.08	0.067	0.20
Cohorts 1, 2, and 3											
Student academic motivation	506	51.0	†	276	50.7	†	230	0.37	1.14	0.746	0.04
Teacher collaboration with and view of colleagues	495	52.1	†	269	46.2	†	226	5.95 ***	1.06	0.000	0.63
Teacher use of data	500	51.8	†	272	48.2	†	228	3.62 ***	0.89	0.000	0.36
Postsecondary educational expectations	510	51.4	†	278	49.0	†	232	2.32 *	0.94	0.014	0.25
View of the school's supports	495	52.1	†	269	48.2	†	226	3.93 ***	0.96	0.000	0.41
Teacher self-efficacy	496	49.5	†	271	49.6	†	225	-0.13	0.54	0.808	-0.01
Strength-based mindset	512	51.1	†	279	49.0	†	233	2.03 **	0.75	0.006	0.21
Students' actual behavior	503	50.4	†	274	48.1	†	229	2.31 **	0.69	0.001	0.23
Perception of students' behavior	505	51.4	†	276	49.3	†	229	2.10 *	0.85	0.014	0.22

Note: SD = standard deviation; SE = standard error. * = statistically significant at the $p < .05$ *level;* ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates. ‡ Survey measure included only in Cohort 3 to better understand technology use during spring 2020 school campus closures due to COVID-19.
Source: American Institutes for Research (AIR) calculations from AIR-administered teacher surveys. Scales are standardized to a mean of 50 and a standard deviation of 10.

Exhibit D2. Impacts on Teacher Experiences Measured by Teacher Surveys for Cohort 1, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Student academic motivation	146	49.7	9.7	58.0	47.4	9.2	88.0	2.38	2.13	0.264	0.25
Teacher collaboration with and view of colleagues	143	57.4	9.0	58.0	45.9	9.4	85.0	11.43 ***	2.55	0.000	1.23
Teacher use of data	144	54.4	9.6	58.0	50.5	10.4	86.0	3.83 *	1.54	0.015	0.38
Postsecondary educational expectations	146	51.6	8.3	58.0	46.1	7.8	88.0	5.41 ***	1.30	0.000	0.67
View of the school's supports	143	51.6	8.3	58.0	50.1	11.1	85.0	1.45	2.51	0.565	0.14
Teacher self-efficacy	144	54.0	8.2	58.0	52.9	9.4	86.0	1.09	1.19	0.362	0.12
Strength-based mindset	147	52.2	7.1	58.0	49.3	9.5	89.0	2.85 *	1.21	0.020	0.33
Students' actual behavior	145	49.3	8.5	58.0	47.2	9.7	87.0	2.12	1.53	0.166	0.23
Perception of students' behavior	146	49.2	9.0	58.00	46.5	9.24	88.00	2.63	1.51	0.083	0.29

Note: SD = standard deviation; SE = standard error. * = statistically significant at the p < .05 level; ** = statistically significant at the p < .01 level; *** = statistically significant at the p < .001 level.

Source: American Institutes for Research (AIR) calculations from AIR-administered teacher surveys. Scales are standardized to a mean of 50 and a standard deviation of 10.

Appendix E. Student Academic, Achievement, and Behavior Outcomes

Exhibits E1 through E24 present the detailed student academic, achievement, and behavior impact analysis results by outcome. For each set of results, we present a table with the main intent-to-treat analysis results followed by a table with the results from a treatment-on-treated analysis using complier average causal effect (CACE) analysis that accounts for the reality that some students whose school was assigned to BARR did not receive the BARR intervention.

Exhibit E1. Impacts on Student Educational Attainment (i.e., Percentage of Core Credits Earned), by Cohort and Subgroup

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
					Cohort 1						
Full Sample	6,003	86.7	26.4	4,060	84.8	26.1	1,943	1.90	2.86	0.508	0.07
Subgroup											
Female	2,856	89.8	21.6	1,925	90.4	23.0	931	-0.60	3.05	0.845	-0.03
Male	3,037	84.9	28.6	2,025	79.9	28.4	1,012	4.94	2.73	0.071	0.17
Students of Color	3,269	83.7	28.5	2,135	81.8	28.7	1,134	1.88	3.95	0.634	0.07
White	2,473	91.6	20.8	1,664	90.8	20.8	809	0.82	2.69	0.759	0.04
FRPL	2,680	87.3	24.4	2,047	81.1	29.1	633	6.19	4.57	0.176	0.24
Not FRPL	1,855	94.2	17.6	1,229	88.9	24.2	626	5.34	2.75	0.052	0.27
ELs	408	80.3	28.1	245	72.5	33.9	163	7.87 ***	1.75	0.000	0.26
Non-ELs	5,028	87.1	25.9	3,251	87.1	24.6	1,777	0.04	2.77	0.989	0.00
SPED	818	86.4	27.4	514	77.6	29.6	304	8.79 **	2.94	0.003	0.31
Non-SPED	4,590	87.6	24.5	2,954	88.6	25.2	1,636	-0.97	3.17	0.761	-0.04
					Cohort 2						
Full Sample	6,822	84.6	25.5	3,332	79.4	31.3	3,490	5.22 *	2.09	0.013	0.18
Subgroup											
Female	3,285	88.0	22.8	1,590	83.5	28.6	1,695	4.46 *	1.81	0.014	0.17
Male	3,527	81.6	27.4	1,736	75.6	33.1	1,791	6.01 *	2.43	0.013	0.20
Students of Color	5,311	82.7	26.9	2,692	76.4	32.7	2,619	6.31 **	2.37	0.008	0.21
White	1,489	93.7	15.2	631	88.5	23.9	858	5.27	2.87	0.066	0.25
FRPL	4,149	83.1	27.0	2,141	74.1	32.6	2,008	9.03 ***	2.74	0.001	0.30
Not FRPL	1,724	91.9	20.2	975	89.7	23.5	749	2.22	1.35	0.100	0.10
ELs	1,732	77.4	31.2	741	72.3	34.2	991	5.10	3.52	0.148	0.15
Non-ELs	5,080	87.7	23.0	2,585	81.3	29.6	2,495	6.43 **	1.99	0.001	0.24
SPED	1,011	81.1	27.2	482	73.4	32.9	529	7.64 *	3.19	0.017	0.25
Non-SPED	5,800	85.3	25.1	2,843	80.5	30.8	2,957	4.83 *	2.03	0.017	0.17

(continues)

Exhibit E1. Impacts on Student Educational Attainment (i.e., Percentage of Core Credits Earned), by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 3											
Full Sample	6,703	91.0	23.7	4,098	87.7	25.4	2,605	3.39 **	1.21	0.005	0.14
Subgroup											
Female	3,239	94.0	18.9	1,999	90.4	22.6	1,240	3.61 **	1.32	0.006	0.18
Male	3,463	88.3	27.1	2,098	85.1	27.5	1,365	3.19 **	1.20	0.008	0.12
Students of Color	4,121	88.2	26.5	2,522	84.6	28.0	1,599	3.64 ***	1.09	0.001	0.13
White	2,575	95.2	17.2	1,574	92.9	19.8	1,001	2.30	1.34	0.087	0.13
FRPL	3,167	87.5	26.6	2,038	83.6	28.0	1,129	3.85	2.02	0.057	0.14
Not FRPL	1,896	96.9	14.2	796	94.9	17.0	1,100	1.97 *	0.81	0.015	0.12
ELs	769	78.0	31.7	563	76.4	33.3	206	1.57	2.36	0.505	0.05
Non-ELs	5,921	92.9	21.6	3,532	89.1	24.0	2,389	3.77 ***	1.14	0.001	0.17
SPED	975	83.4	29.4	613	81.0	30.7	362	2.48	2.23	0.265	0.08
Non-SPED	5,715	92.4	22.4	3,482	88.8	23.9	2,233	3.58 **	1.19	0.003	0.16
Cohorts 1, 2, and 3											
Full Sample	19,528	89.1	†	11,490	85.5	†	8,038	3.62 ***	0.98	0.000	0.14
Subgroup											
Female	9,380	91.7	†	5,514	88.3	†	3,866	3.42 ***	1.00	0.001	0.15
Male	10,027	86.7	†	5,859	82.8	†	4,168	3.91 ***	1.00	0.000	0.14
Students of Color	12,701	87.1	†	7,349	83.1	†	5,352	3.98 ***	0.96	0.000	0.14
White	6,537	94.4	†	3,869	91.9	†	2,668	2.49 *	1.11	0.024	0.13
FRPL	9,996	86.1	†	6,226	80.3	†	3,770	5.74 ***	1.53	0.000	0.21
Not FRPL	5,475	95.5	†	3,000	93.2	†	2,475	2.24 ***	0.67	0.001	0.13
ELs	2,909	79.2	†	1,549	73.7	†	1,360	5.56 ***	1.31	0.000	0.18
Non-ELs	16,029	91.1	†	9,368	87.2	†	6,661	3.94 ***	0.93	0.000	0.17
SPED	2,804	83.7	†	1,609	78.3	†	1,195	5.45 ***	1.55	0.000	0.18
Non-SPED	16,105	90.3	†	9,279	86.9	†	6,826	3.44 ***	0.98	0.000	0.14

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: American Institutes for Research (AIR) calculations from school-provided administrative data.

Exhibit E2. Impacts on Student Educational Attainment (i.e., Percentage of Core Credits Earned), by Cohort and Subgroup, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	6,003	88.3	25.1	1,892	85.0	26.8	4,111	3.32	5.51	0.547	0.13
Subgroup											
Female	2,856	89.3	21.8	956	90.3	22.2	1,900	-1.02	5.06	0.841	-0.05
Male	3,037	89.3	27.8	935	80.5	28.9	2,102	8.79	6.44	0.172	0.31
Students of Color	3,269	88.7	29.6	837	81.0	28.2	2,432	7.66	19.92	0.701	0.27
White	2,473	91.9	20.3	1,050	90.9	21.2	1,423	1.00	3.30	0.762	0.05
FRPL	2,680	91.4	24.5	936	82.8	26.2	1,744	8.63	7.37	0.241	0.34
Not FRPL	1,855	96.2	17.7	754	89.8	21.6	1,101	6.45	3.64	0.076	0.32
ELs	408	110.0	21.1	77	69.6	32.2	331	40.38	32.94	0.221	1.33
Non-ELs	5,028	87.2	25.4	1,658	87.1	25.5	3,370	0.06	4.31	0.989	0.00
SPED	818	96.2	29.8	196	79.0	27.8	622	17.18	9.10	0.059	0.61
Non-SPED	4,590	87.0	24.1	1,611	88.5	25.0	2,979	-1.50	4.71	0.751	-0.06
Cohort 2											
Full Sample	6,822	87.6	24.7	1,802	79.9	29.9	5,020	7.66 **	2.92	0.009	0.27
Subgroup											
Female	3,285	90.7	20.9	852	83.9	27.4	2,433	6.76 *	2.67	0.011	0.26
Male	3,527	84.8	27.3	950	76.3	31.6	2,577	8.51 **	3.21	0.008	0.28
Students of Color	5,311	86.4	27.3	1,272	77.5	30.9	4,039	8.87 **	3.21	0.006	0.30
White	1,489	96.3	14.3	527	87.6	23.5	962	8.63 ***	2.59	0.001	0.42
FRPL	4,149	88.8	26.6	945	75.7	30.9	3,204	13.08 ***	3.42	0.000	0.44
Not FRPL	1,724	93.4	20.2	647	89.4	22.6	1,077	3.96	2.37	0.095	0.18
ELs	1,732	78.9	32.3	515	72.6	33.3	1,217	6.33	4.66	0.175	0.19
Non-ELs	5,080	92.4	19.4	1,287	81.9	28.4	3,793	10.43 ***	2.29	0.000	0.40
SPED	1,011	86.1	24.6	263	73.9	31.9	748	12.22 **	4.54	0.007	0.40
Non-SPED	5,800	88.0	24.7	1,538	81.0	29.4	4,262	6.99 *	2.87	0.015	0.25

(continues)

Exhibit E2. Impacts on Student Educational Attainment (i.e., Percentage of Core Credits Earned), by Cohort and Subgroup, *Compliance Adjusted* (continued)

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
					Cohort 3						
Full Sample	6,703	91.4	24.3	3,810	87.6	24.5	2,893	3.80 **	1.33	0.004	0.16
Subgroup											
Female	3,239	94.3	19.4	1,866	90.3	21.8	1,373	4.01 **	1.45	0.006	0.20
Male	3,463	88.6	27.8	1,943	85.0	26.6	1,520	3.60 **	1.31	0.006	0.13
Students of Color	4,121	88.6	27.0	2,352	84.5	27.2	1,769	4.08 **	1.29	0.002	0.15
White	2,575	95.5	17.9	1,456	92.9	18.8	1,119	2.63	1.48	0.076	0.14
FRPL	3,167	87.6	26.7	1,989	83.6	27.8	1,178	3.94	2.06	0.056	0.15
Not FRPL	1,896	97.2	14.6	708	94.9	16.7	1,188	2.29 *	0.94	0.014	0.14
ELs	769	78.0	31.8	546	76.5	33.1	223	1.58	2.34	0.499	0.05
Non-ELs	5,921	93.3	22.2	3,261	89.0	23.1	2,660	4.28 ***	1.24	0.001	0.19
SPED	975	83.5	29.6	591	81.0	30.4	384	2.54	2.28	0.264	0.08
Non-SPED	5,715	92.8	23.0	3,216	88.7	22.9	2,499	4.06 **	1.27	0.001	0.18
					Cohorts 1, 2, and 3						
Full Sample	19,528	90.6	†	7,504	86.2	†	12,024	4.41 ***	1.18	0.000	0.18
Subgroup											
Female	9,380	93.3	†	3,674	89.0	†	5,706	4.30 ***	1.23	0.000	0.20
Male	10,027	88.1	†	3,828	83.6	†	6,199	4.45 ***	1.19	0.000	0.16
Students of Color	12,701	88.3	†	4,461	83.5	†	8,240	4.76 ***	1.20	0.000	0.17
White	6,537	95.2	†	3,033	91.5	†	3,504	3.70 **	1.20	0.002	0.19
FRPL	9,996	88.1	†	3,870	81.6	†	6,126	6.50 ***	1.72	0.000	0.23
Not FRPL	5,475	96.6	†	2,109	93.9	†	3,366	2.73 **	0.85	0.001	0.16
ELs	2,909	78.3	†	1,138	75.7	†	1,771	2.68	2.08	0.198	0.08
Non-ELs	16,029	92.7	†	6,206	87.4	†	9,823	5.33 ***	1.05	0.000	0.23
SPED	2,804	84.6	†	1,050	79.5	†	1,754	5.09 *	1.99	0.010	0.17
Non-SPED	16,105	91.7	†	6,365	87.5	†	9,740	4.20 ***	1.13	0.000	0.18

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E3. Impacts on Course Failure (i.e., Percentage of Students Who Passed All Their Core Courses), by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	6,003	73.4	46.2	4,060	71.3	45.9	1,943	2.09	31.53	0.741	0.06
Subgroup											
Female	2,856	79.2	42.7	1,925	79.9	44.1	931	-0.75	37.02	0.901	-0.03
Male	3,037	69.8	47.9	2,025	63.5	47.3	1,012	6.36	27.06	0.289	0.17
Students of Color	3,269	66.7	47.9	2,135	64.6	48.6	1,134	2.02	30.20	0.767	0.05
White	2,473	82.7	41.3	1,664	82.5	39.5	809	0.24	45.68	0.971	0.01
FRPL	2,680	71.7	46.3	2,047	60.1	48.5	633	11.60	37.43	0.164	0.32
Not FRPL	1,855	88.3	37.0	1,229	78.9	40.8	626	9.35	35.88	0.052	0.42
ELs	408	63.1	49.3	245	46.4	49.7	163	16.74 **	24.07	0.005	0.41
Non-ELs	5,028	75.1	45.4	3,251	74.9	44.8	1,777	0.21	31.94	0.972	0.01
SPED	818	67.0	48.1	514	55.3	48.8	304	11.68	31.15	0.113	0.30
Non-SPED	4,590	77.7	43.9	2,954	78.7	45.2	1,636	-0.98	35.49	0.871	-0.03
Cohort 2											
Full Sample	6,822	67.3	47.4	3,332	58.9	49.5	3,490	8.43 *	17.12	0.034	0.22
Subgroup											
Female	3,285	74.0	45.2	1,590	67.1	48.2	1,695	6.96 *	17.10	0.050	0.20
Male	3,527	61.5	48.7	1,736	51.4	50.0	1,791	10.06 *	17.44	0.019	0.25
Students of Color	5,311	63.7	48.4	2,692	52.9	50.0	2,619	10.78 **	16.91	0.008	0.27
White	1,489	84.2	38.6	631	80.8	45.3	858	3.43	29.81	0.423	0.14
FRPL	4,149	62.1	49.0	2,141	47.6	50.0	2,008	14.52 **	18.68	0.002	0.36
Not FRPL	1,724	84.6	40.8	975	78.6	42.6	749	6.06 *	19.37	0.036	0.25
ELs	1,732	53.7	49.9	741	46.3	50.0	991	7.43	25.33	0.240	0.18
Non-ELs	5,080	72.6	46.0	2,585	62.3	48.9	2,495	10.25 *	18.64	0.012	0.28
SPED	1,011	58.4	49.2	482	47.5	49.9	529	10.96	26.50	0.096	0.27
Non-SPED	5,800	69.8	46.9	2,843	61.4	49.3	2,957	8.44 *	16.88	0.026	0.23

(continues)

Exhibit E3. Impacts on Course Failure (i.e., Percentage of Students Who Passed All Their Core Courses), by Cohort and Subgroup (continued)

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
					Cohort 3						
Full Sample	6,703	87.8	40.2	4,098	82.8	40.1	2,605	4.98 **	12.30	0.001	0.24
Subgroup											
Female	3,239	91.0	35.9	1,999	87.0	37.5	1,240	3.97 **	12.55	0.001	0.25
Male	3,463	85.0	43.3	2,098	79.2	42.1	1,365	5.78 **	13.23	0.003	0.24
Students of Color	4,121	83.8	43.7	2,522	77.9	43.3	1,599	5.84 ***	10.18	0.000	0.23
White	2,575	93.9	31.8	1,574	91.2	32.8	1,001	2.66	28.37	0.169	0.24
FRPL	3,167	78.4	43.5	2,038	69.7	46.8	1,129	8.65 *	22.83	0.047	0.27
Not FRPL	1,896	93.4	28.9	796	92.3	29.1	1,100	1.13	26.89	0.525	0.10
ELs	769	66.9	49.3	563	61.6	50.1	206	5.27	17.34	0.185	0.14
Non-ELs	5,921	90.6	37.4	3,532	85.8	38.0	2,389	4.81 ***	12.81	0.000	0.28
SPED	975	72.3	46.4	613	66.4	48.4	362	5.92	19.11	0.144	0.17
Non-SPED	5,715	91.0	38.6	3,482	86.5	37.7	2,233	4.49 ***	11.68	0.000	0.28
					Cohorts 1, 2, and 3						
Full Sample	19,528	80.2	†	11,490	74.4	†	8,038	5.78 ***	9.52	0.000	0.21
Subgroup											
Female	9,380	84.6	†	5,514	80.0	†	3,866	4.62 ***	9.76	0.000	0.20
Male	10,027	75.5	†	5,859	68.3	†	4,168	7.22 ***	9.82	0.000	0.23
Students of Color	12,701	77.5	†	7,349	70.8	†	5,352	6.76 ***	8.38	0.000	0.23
White	6,537	88.2	†	3,869	85.6	†	2,668	2.56	18.74	0.153	0.14
FRPL	9,996	69.0	†	6,226	56.9	†	3,770	12.10 ***	13.48	0.000	0.33
Not FRPL	5,475	87.7	†	3,000	82.6	†	2,475	5.18 **	14.39	0.007	0.26
ELs	2,909	62.8	†	1,549	54.0	†	1,360	8.78 **	12.30	0.003	0.22
Non-ELs	16,029	83.9	†	9,368	78.0	†	6,661	5.93 ***	10.03	0.000	0.25
SPED	2,804	67.5	†	1,609	59.0	†	1,195	8.45 **	13.88	0.008	0.23
Non-SPED	16,105	83.7	†	9,279	78.4	†	6,826	5.31 ***	9.27	0.000	0.23

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E4. Impacts on Course Failure (i.e., Percentage of Students Who Passed All Their Core Courses), by Cohort and Subgroup, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	6,003	75.9	43.5	1,892	72.4	47.0	4,111	3.49	55.18	0.764	0.11
Subgroup											
Female	2,856	79.0	40.6	956	80.2	44.3	1,900	-1.28	63.13	0.826	-0.05
Male	3,037	75.4	45.8	935	64.8	48.3	2,102	10.62	48.18	0.370	0.31
Students of Color	3,269	72.3	46.3	837	64.5	48.6	2,432	7.85	122.93	0.861	0.22
White	2,473	83.0	40.6	1,050	82.7	40.8	1,423	0.29	55.58	0.927	0.01
FRPL	2,680	78.3	44.8	936	63.6	47.8	1,744	14.73	52.25	0.241	0.44
Not FRPL	1,855	90.8	36.2	754	80.9	39.7	1,101	9.86	43.38	0.103	0.51
ELs	408	95.5	42.6	77	39.3	50.0	331	56.27	123.49	0.155	2.12
Non-ELs	5,028	75.7	42.9	1,658	75.4	46.1	3,370	0.33	49.65	0.995	0.01
SPED	818	78.5	49.2	196	58.1	48.1	622	20.36	60.90	0.178	0.59
Non-SPED	4,590	77.2	41.8	1,611	78.8	45.5	2,979	-1.53	54.94	0.876	-0.05
Cohort 2											
Full Sample	6,822	72.1	45.5	1,802	60.3	49.4	5,020	11.83 *	25.12	0.028	0.32
Subgroup											
Female	3,285	77.9	42.4	852	68.0	48.0	2,433	9.95	25.88	0.063	0.31
Male	3,527	67.0	47.6	950	53.2	50.0	2,577	13.84 *	24.70	0.011	0.35
Students of Color	5,311	69.5	47.4	1,272	54.9	49.8	4,039	14.59 **	23.79	0.010	0.38
White	1,489	85.7	38.2	527	80.2	44.9	962	5.48	48.78	0.237	0.24
FRPL	4,149	70.6	47.8	945	50.4	50.0	3,204	20.11 ***	27.06	0.000	0.52
Not FRPL	1,724	88.6	40.0	647	79.0	42.5	1,077	9.61	34.60	0.079	0.44
ELs	1,732	57.0	49.7	515	47.8	50.0	1,217	9.19	31.43	0.257	0.22
Non-ELs	5,080	79.0	42.4	1,287	63.8	48.7	3,793	15.26 ***	30.22	0.001	0.46
SPED	1,011	65.1	46.9	263	48.0	50.0	748	17.19	42.42	0.079	0.43
Non-SPED	5,800	74.5	45.3	1,538	62.9	49.0	4,262	11.59 *	24.44	0.031	0.33

(continues)

Exhibit E4. Impacts on Course Failure (i.e., Percentage of Students Who Passed All Their Core Courses), by Cohort and Subgroup, *Compliance Adjusted* (continued)

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
					Cohort 3						
Full Sample	6,703	88.9	41.1	3,810	83.6	38.8	2,893	5.27 **	13.77	0.003	0.27
Subgroup											
Female	3,239	91.6	36.7	1,866	87.4	36.2	1,373	4.24 ***	13.95	0.000	0.28
Male	3,463	86.0	44.2	1,943	79.8	40.8	1,520	6.27 *	14.90	0.023	0.27
Students of Color	4,121	84.4	44.5	2,352	77.9	42.2	1,769	6.45 ***	11.40	0.000	0.26
White	2,575	94.4	32.8	1,456	91.4	31.4	1,119	2.91	32.51	0.229	0.27
FRPL	3,167	78.5	43.7	1,989	69.7	46.5	1,178	8.83 *	23.34	0.017	0.28
Not FRPL	1,896	95.1	30.2	708	94.1	28.3	1,188	1.01	31.25	0.275	0.12
ELs	769	61.6	49.4	546	56.1	49.9	223	5.58	17.40	0.228	0.14
Non-ELs	5,921	91.4	38.5	3,261	86.2	36.7	2,660	5.21 **	14.53	0.001	0.32
SPED	975	72.2	46.7	591	66.1	48.0	384	6.10	19.59	0.179	0.17
Non-SPED	5,715	91.7	39.6	3,216	86.8	36.3	2,499	4.90 **	13.26	0.001	0.31
					Cohorts 1, 2, and 3						
Full Sample	19,528	84.6	†	7,504	77.9	†	12,024	6.63 ***	11.79	0.000	0.28
Subgroup											
Female	9,380	88.2	†	3,674	82.9	†	5,706	5.28 ***	12.05	0.000	0.27
Male	10,027	80.6	†	3,828	72.2	†	6,199	8.44 ***	12.33	0.000	0.30
Students of Color	12,701	81.5	†	4,461	73.6	†	8,240	7.97 ***	10.25	0.000	0.29
White	6,537	90.0	†	3,033	87.0	†	3,504	3.05	24.32	0.148	0.19
FRPL	9,996	75.5	†	3,870	61.7	†	6,126	13.75 ***	16.74	0.000	0.40
Not FRPL	5,475	91.9	†	2,109	85.9	†	3,366	5.98 *	20.45	0.010	0.39
ELs	2,909	61.1	†	1,138	53.9	†	1,771	7.17 *	15.11	0.039	0.18
Non-ELs	16,029	88.2	†	6,206	81.5	†	9,823	6.66 ***	12.66	0.000	0.33
SPED	2,804	71.5	†	1,050	62.5	†	1,754	9.02 *	17.07	0.017	0.25
Non-SPED	16,105	87.3	†	6,365	81.2	†	9,740	6.08 ***	11.40	0.000	0.30

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E5. Impacts on GPA, by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	5,915	2.48	1.07	4,059	2.58	1.08	1,856	-0.10	0.14	0.450	-0.10
Subgroup											
Female	2,807	2.66	1.02	1,926	2.84	1.07	881	-0.18	0.15	0.241	-0.17
Male	2,998	2.35	1.05	2,023	2.38	1.06	975	-0.03	0.13	0.795	-0.03
Students of Color	3,280	2.35	1.09	2,139	2.37	1.11	1,141	-0.02	0.13	0.856	-0.02
White	2,378	2.74	0.97	1,663	2.94	0.88	715	-0.20	0.17	0.251	-0.21
FRPL	2,626	2.49	1.01	2,033	2.29	1.07	593	0.20	0.21	0.358	0.19
Not FRPL	1,806	2.95	0.89	1,230	2.88	0.95	576	0.08	0.15	0.621	0.08
ELs	407	2.27	1.04	242	1.79	1.08	165	0.48 ***	0.11	0.000	0.45
Non-ELs	4,943	2.51	1.06	3,252	2.67	1.05	1,691	-0.16	0.14	0.246	-0.15
SPED	808	2.14	0.98	509	2.17	1.05	299	-0.03	0.08	0.675	-0.03
Non-SPED	4,514	2.58	1.03	2,957	2.75	1.07	1,557	-0.17	0.17	0.316	-0.16
Cohort 2											
Full Sample	6,859	2.42	0.99	3,288	2.34	1.10	3,571	0.08	0.08	0.321	0.07
Subgroup											
Female	3,283	2.61	0.96	1,562	2.51	1.08	1,721	0.09	0.07	0.216	0.09
Male	3,565	2.25	0.99	1,720	2.19	1.08	1,845	0.07	0.08	0.413	0.06
Students of Color	5,424	2.33	0.98	2,736	2.22	1.07	2,688	0.11	0.08	0.197	0.11
White	1,409	2.83	0.89	542	2.77	1.07	867	0.06	0.15	0.673	0.06
FRPL	4,158	2.27	0.97	2,108	2.10	1.07	2,050	0.17	0.10	0.083	0.17
Not FRPL	1,705	2.82	0.92	951	2.86	1.00	754	-0.03	0.13	0.799	-0.03
ELs	1,798	2.10	0.96	771	2.16	1.02	1,027	-0.06	0.10	0.584	-0.06
Non-ELs	5,050	2.55	0.97	2,511	2.39	1.12	2,539	0.17 *	0.08	0.031	0.16
SPED	1,020	2.08	0.93	471	1.97	1.02	549	0.11	0.09	0.227	0.11
Non-SPED	5,827	2.48	0.99	2,810	2.41	1.10	3,017	0.07	0.08	0.363	0.07

(continues)

Exhibit E5. Impacts on GPA, by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 3											
Full Sample	6,842	2.83	0.98	4,092	2.60	1.13	2,750	0.23 ***	0.07	0.000	0.22
Subgroup											
Female	3,316	3.01	0.89	2,000	2.78	1.09	1,316	0.23 ***	0.06	0.000	0.24
Male	3,525	2.66	1.02	2,091	2.44	1.16	1,434	0.22 **	0.07	0.001	0.21
Students of Color	4,218	2.68	1.00	2,516	2.36	1.18	1,702	0.32 ***	0.05	0.000	0.30
White	2,333	3.01	0.91	1,290	2.99	0.94	1,043	0.02	0.06	0.793	0.02
FRPL	3,203	2.55	0.98	2,041	2.32	1.07	1,162	0.24 **	0.08	0.002	0.23
Not FRPL	1,961	3.24	0.78	796	3.20	0.89	1,165	0.04	0.07	0.543	0.05
ELs	780	2.20	0.94	561	1.88	1.00	219	0.32 ***	0.06	0.000	0.33
Non-ELs	6,046	2.92	0.95	3,527	2.69	1.12	2,519	0.23 ***	0.07	0.001	0.22
SPED	983	2.27	1.02	610	2.06	1.11	373	0.21	0.12	0.076	0.20
Non-SPED	5,843	2.93	0.94	3,478	2.69	1.10	2,365	0.25 ***	0.06	0.000	0.25
Cohorts 1, 2, and 3											
Full Sample	19,616	2.64	†	11,439	2.50	†	8,177	0.13 **	0.05	0.004	0.13
Subgroup											
Female	9,406	2.82	†	5,488	2.68	†	3,918	0.14 **	0.05	0.002	0.14
Male	10,088	2.47	†	5,834	2.34	†	4,254	0.13 **	0.05	0.007	0.12
Students of Color	12,922	2.56	†	7,391	2.33	†	5,531	0.24 ***	0.04	0.000	0.22
White	6,120	2.96	†	3,495	2.96	†	2,625	0.00	0.05	0.973	0.00
FRPL	9,987	2.45	†	6,182	2.24	†	3,805	0.21 ***	0.06	0.000	0.21
Not FRPL	5,472	3.12	†	2,977	3.08	†	2,495	0.03	0.06	0.573	0.04
ELs	2,985	2.19	†	1,574	1.92	†	1,411	0.27 ***	0.05	0.000	0.28
Non-ELs	16,039	2.73	†	9,290	2.57	†	6,749	0.16 ***	0.05	0.001	0.15
SPED	2,811	2.14	†	1,590	2.08	†	1,221	0.07	0.05	0.219	0.06
Non-SPED	16,184	2.74	†	9,245	2.59	†	6,939	0.15 **	0.05	0.001	0.14

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E6. Impacts on GPA, by Cohort and Subgroup, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	5,915	2.38	1.03	1,901	2.57	1.08	4,014	-0.19	0.22	0.405	-0.18
Subgroup											
Female	2,807	2.51	0.99	958	2.83	1.05	1,849	-0.31	0.23	0.179	-0.30
Male	2,998	2.32	1.03	942	2.38	1.06	2,056	-0.06	0.23	0.787	-0.06
Students of Color	3,280	2.28	1.10	844	2.38	1.09	2,436	-0.10	0.51	0.845	-0.09
White	2,378	2.67	0.95	1,052	2.91	0.95	1,326	-0.24	0.20	0.237	-0.26
FRPL	2,626	2.62	1.02	937	2.34	1.01	1,689	0.28	0.33	0.402	0.27
Not FRPL	1,806	2.98	0.88	756	2.89	0.93	1,050	0.09	0.19	0.626	0.10
ELs	407	4.04	0.89	77	1.61	1.08	330	2.43	2.38	0.307	2.32
Non-ELs	4,943	2.40	1.04	1,666	2.65	1.06	3,277	-0.25	0.19	0.191	-0.24
SPED	808	2.10	0.93	202	2.17	1.03	606	-0.07	0.15	0.661	-0.07
Non-SPED	4,514	2.46	1.02	1,613	2.73	1.05	2,901	-0.27	0.23	0.253	-0.26
Cohort 2											
Full Sample	6,859	2.47	0.99	1,745	2.35	1.06	5,114	0.12	0.13	0.363	0.12
Subgroup											
Female	3,283	2.67	0.95	815	2.52	1.05	2,468	0.15	0.14	0.275	0.15
Male	3,565	2.29	1.00	930	2.19	1.05	2,635	0.10	0.13	0.438	0.10
Students of Color	5,424	2.40	0.99	1,303	2.24	1.04	4,121	0.16	0.14	0.252	0.16
White	1,409	2.91	0.89	438	2.73	1.06	971	0.18	0.43	0.679	0.18
FRPL	4,158	2.40	0.99	902	2.13	1.03	3,256	0.27	0.18	0.135	0.27
Not FRPL	1,705	2.80	0.92	620	2.86	0.98	1,085	-0.06	0.23	0.792	-0.06
ELs	1,798	2.09	0.97	536	2.16	1.01	1,262	-0.07	0.12	0.565	-0.07
Non-ELs	5,050	2.70	0.95	1,209	2.40	1.07	3,841	0.31	0.16	0.061	0.29
SPED	1,020	2.16	0.89	247	1.97	1.00	773	0.19	0.17	0.254	0.20
Non-SPED	5,827	2.52	1.00	1,497	2.42	1.06	4,330	0.11	0.13	0.402	0.10

(continues)

Exhibit E6. Impacts on GPA, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 3											
Full Sample	6,842	2.85	0.98	3,804	2.60	1.11	3,038	0.25 ***	0.07	0.001	0.24
Subgroup											
Female	3,316	3.03	0.90	1,867	2.78	1.06	1,449	0.26 ***	0.07	0.000	0.26
Male	3,525	2.68	1.02	1,936	2.43	1.14	1,589	0.25 **	0.08	0.001	0.23
Students of Color	4,218	2.71	1.00	2,346	2.35	1.17	1,872	0.36 ***	0.06	0.000	0.33
White	2,333	3.01	0.92	1,172	2.99	0.91	1,161	0.02	0.07	0.795	0.02
FRPL	3,203	2.56	0.98	1,992	2.32	1.07	1,211	0.24 **	0.08	0.002	0.24
Not FRPL	1,961	3.25	0.79	708	3.20	0.88	1,253	0.05	0.08	0.548	0.06
ELs	780	2.21	0.94	544	1.89	1.00	236	0.32 ***	0.06	0.000	0.33
Non-ELs	6,046	2.94	0.96	3,256	2.68	1.09	2,790	0.26 ***	0.08	0.001	0.25
SPED	983	2.28	1.03	588	2.06	1.11	395	0.22	0.12	0.075	0.20
Non-SPED	5,843	2.96	0.95	3,212	2.68	1.07	2,631	0.28 ***	0.07	0.000	0.28
Cohorts 1, 2, and 3											
Full Sample	19,616	2.73	†	7,450	2.54	†	12,166	0.19 **	0.06	0.002	0.18
Subgroup											
Female	9,406	2.93	†	3,640	2.73	†	5,766	0.20 ***	0.06	0.001	0.20
Male	10,088	2.56	†	3,808	2.37	†	6,280	0.19 **	0.06	0.003	0.18
Students of Color	12,922	2.66	†	4,493	2.33	†	8,429	0.33 ***	0.05	0.000	0.30
White	6,120	2.97	†	2,662	2.98	†	3,458	0.00	0.06	0.955	0.00
FRPL	9,987	2.5	†	3,831	2.3	†	6,156	0.25 ***	0.07	0.000	0.24
Not FRPL	5,472	3.2	†	2,084	3.1	†	3,388	0.04	0.07	0.533	0.05
ELs	2,985	2.18	†	1,157	1.94	†	1,828	0.24 ***	0.06	0.000	0.25
Non-ELs	16,039	2.84	†	6,131	2.63	†	9,908	0.21 **	0.07	0.001	0.20
SPED	2,811	2.20	†	1,037	2.07	†	1,774	0.13	0.08	0.124	0.12
Non-SPED	16,184	2.83	†	6,322	2.63	†	9,862	0.21 ***	0.06	0.001	0.20

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E7. Impacts on PSAT Reading Scores, by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	3,443	433	89	2,193	422	84	1,250	11.0	5.71	0.055	0.13
Subgroup											
Female	1,712	438	88	1,101	436	84	611	2.4	4.36	0.590	0.03
Male	1,731	428	90	1,092	408	82	639	19.9 *	9.36	0.034	0.23
Students of Color	1,652	420	89	1,047	397	72	605	23.1 ***	5.89	0.000	0.28
White	1,789	445	88	1,144	445	87	645	0.5	2.89	0.870	0.01
FRPL	1,417	430	87	952	403	79	465	27.1 ***	7.30	0.000	0.32
Not FRPL	1,460	454	89	921	448	87	539	5.9	5.45	0.280	0.07
ELs	173	368	65	95	338	39	78	29.3 **	9.70	0.003	0.53
Non-ELs	2,928	436	88	1,756	426	84	1,172	10.1	5.62	0.073	0.12
SPED	321	370	75	186	367	58	135	3.5	7.75	0.649	0.05
Non-SPED	2,834	443	88	1,719	435	83	1,115	7.3	6.13	0.232	0.08
Cohort 2											
Full Sample	3,829	425	90	1,777	422	88	2,052	2.3	4.82	0.631	0.03
Subgroup											
Female	1,959	431	88	894	434	88	1,065	-2.4	4.75	0.612	-0.03
Male	1,870	418	92	883	411	87	987	7.3	4.93	0.138	0.08
Students of Color	2,710	404	83	1,270	399	73	1,440	4.6	4.85	0.345	0.06
White	1,113	471	90	506	484	93	607	-13.1 ***	2.98	0.000	-0.14
FRPL	2,018	402	76	940	396	74	1,078	5.4	3.70	0.141	0.07
Not FRPL	1,252	473	95	701	474	94	551	-1.2	6.92	0.862	-0.01
ELs	918	375	59	345	364	56	573	10.7 **	3.35	0.001	0.19
Non-ELs	2,911	441	91	1,432	441	89	1,479	-0.6	5.16	0.910	-0.01
SPED	412	361	72	208	362	62	204	-0.5	8.40	0.956	-0.01
Non-SPED	3,416	433	89	1,568	429	88	1,848	3.6	4.67	0.438	0.04

(continues)

Exhibit E7. Impacts on PSAT Reading Scores, by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	7,272	428	†	3,970	422	†	3,302	5.9	3.68	0.109	0.07
Subgroup											
Female	3,671	435	†	1,995	435	†	1,676	0.2	3.21	0.957	0.00
Male	3,601	420	†	1,975	410	†	1,626	10.1 *	4.37	0.021	0.11
Students of Color	4,362	410	†	2,317	398	†	2,045	12.1 **	3.74	0.001	0.15
White	2,902	458	†	1,650	464	†	1,252	-6.1 **	2.07	0.003	-0.07
FRPL	3,435	408	†	1,892	398	†	1,543	9.9 **	3.30	0.003	0.13
Not FRPL	2,712	461	†	1,622	458	†	1,090	3.2	4.28	0.458	0.03
ELs	1,091	374	†	440	361	†	651	12.7 ***	3.17	0.000	0.22
Non-ELs	5,839	439	†	3,188	434	†	2,651	4.3	3.80	0.258	0.05
SPED	733	366	†	394	364	†	339	1.7	5.69	0.766	0.02
Non-SPED	6,250	437	†	3,287	432	†	2,963	5.0	3.71	0.180	0.06

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E8. Impacts on PSAT Reading Scores, by Cohort and Subgroup, *Compliance Adjusted*

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
					Cohort 1						
Full Sample	3,443	437	88	1,370	424	86	2,073	13.4	7.25	0.065	0.15
Subgroup											
Female	1,712	439	89	711	436	85	1,001	2.9	5.26	0.588	0.03
Male	1,731	436	88	659	412	86	1,072	24.3 *	12.31	0.049	0.28
Students of Color	1,652	431	88	546	402	80	1,106	29.0 ***	7.28	0.000	0.35
White	1,789	446	88	823	445	87	966	0.5	3.31	0.869	0.01
FRPL	1,417	439	87	677	405	82	740	34.1 ***	9.77	0.000	0.40
Not FRPL	1,460	456	89	633	449	89	827	7.0	6.45	0.281	0.08
ELs	173	382	63	50	343	51	123	38.4 *	16.37	0.020	0.70
Non-ELs	2,928	439	88	1,198	427	84	1,730	12.1	6.87	0.079	0.14
SPED	321	372	67	98	368	70	223	3.9	8.44	0.645	0.06
Non-SPED	2,834	445	88	1,220	436	85	1,614	8.7	7.40	0.238	0.10
					Cohort 2						
Full Sample	3,829	426	86	1,316	422	91	2,513	3.1	6.32	0.619	0.04
Subgroup											
Female	1,959	430	83	652	434	90	1,307	-3.3	6.71	0.622	-0.04
Male	1,870	421	88	664	411	91	1,206	9.9	5.99	0.100	0.11
Students of Color	2,710	405	73	885	399	80	1,825	5.9	5.79	0.310	0.08
White	1,113	463	91	430	487	92	683	-24.2 **	7.90	0.002	-0.26
FRPL	2,018	404	74	681	396	76	1,337	7.1	4.49	0.114	0.09
Not FRPL	1,252	472	91	499	474	97	753	-2.1	12.37	0.864	-0.02
ELs	918	377	61	285	364	55	633	12.9 **	4.35	0.003	0.23
Non-ELs	2,911	441	88	1,031	441	92	1,880	-0.9	7.66	0.911	-0.01
SPED	412	361	73	161	362	63	251	-0.6	11.46	0.955	-0.01
Non-SPED	3,416	434	85	1,154	429	91	2,262	4.9	5.93	0.404	0.06

(continues)

Exhibit E8. Impacts on PSAT Reading Scores, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	7,272	431	†	2,686	423	†	4,586	7.6	4.76	0.112	0.09
Subgroup											
Female	3,671	436	†	1,363	435	†	2,308	0.5	4.14	0.902	0.01
Male	3,601	423	†	1,323	411	†	2,278	12.6 *	5.38	0.019	0.14
Students of Color	4,362	415	†	1,431	401	†	2,931	14.8 **	4.53	0.001	0.19
White	2,902	448	†	1,253	451	†	1,649	-3.2	3.05	0.302	-0.04
FRPL	3,435	410	†	1,358	398	†	2,077	11.8 **	4.08	0.004	0.15
Not FRPL	2,712	459	†	1,132	454	†	1,580	5.0	5.72	0.380	0.06
ELs	1,091	377	†	335	363	†	756	14.6 ***	4.20	0.001	0.26
Non-ELs	5,839	440	†	2,229	434	†	3,610	6.3	5.12	0.217	0.07
SPED	733	368	†	259	366	†	474	2.3	6.80	0.736	0.03
Non-SPED	6,250	438	†	2,374	432	†	3,876	6.4	4.63	0.165	0.07

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E9. Impacts on PSAT Math Scores, by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	3,443	421	80	2,193	410	74	1,250	10.7 **	3.36	0.001	0.14
Subgroup											
Female	1,712	419	79	1,101	410	72	611	8.8 *	3.90	0.025	0.11
Male	1,731	422	80	1,092	410	76	639	12.8 *	5.83	0.029	0.16
Students of Color	1,652	406	80	1,047	390	65	605	15.5 **	5.47	0.005	0.21
White	1,789	435	78	1,144	427	75	645	8.7 **	3.07	0.005	0.11
FRPL	1,417	416	75	952	396	65	465	20.2 ***	5.31	0.000	0.28
Not FRPL	1,460	442	82	921	430	76	539	12.1 **	4.50	0.007	0.15
ELs	173	365	56	95	347	55	78	17.6	9.84	0.076	0.32
Non-ELs	2,928	422	78	1,756	410	74	1,172	11.2 **	3.82	0.003	0.15
SPED	321	364	72	186	370	64	135	-5.2	3.14	0.101	-0.08
Non-SPED	2,834	429	79	1,719	420	73	1,115	8.8 *	4.06	0.030	0.11
Cohort 2											
Full Sample	3,829	424	73	1,777	423	73	2,052	1.0	3.16	0.761	0.01
Subgroup											
Female	1,959	420	68	894	423	66	1,065	-2.8	3.46	0.421	-0.04
Male	1,870	427	78	883	423	79	987	4.4	3.65	0.225	0.06
Students of Color	2,710	412	68	1,270	407	63	1,440	5.3 *	2.67	0.050	0.08
White	1,113	444	77	506	468	80	607	-23.4 ***	5.21	0.000	-0.30
FRPL	2,018	408	65	940	403	63	1,078	4.8 *	2.18	0.027	0.08
Not FRPL	1,252	454	79	701	462	79	551	-7.8	4.42	0.076	-0.10
ELs	918	398	54	345	396	57	573	2.7	4.50	0.554	0.05
Non-ELs	2,911	431	76	1,432	432	75	1,479	-0.7	3.67	0.855	-0.01
SPED	412	375	57	208	378	57	204	-3.4	3.19	0.289	-0.06
Non-SPED	3,416	430	73	1,568	428	72	1,848	1.9	3.15	0.542	0.03

(continues)

Exhibit E9. Impacts on PSAT Math Scores, by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	7,272	422	†	3,970	417	†	3,302	5.5 *	2.30	0.016	0.07
Subgroup											
Female	3,671	420	†	1,995	417	†	1,676	2.3	2.59	0.376	0.03
Male	3,601	426	†	1,975	419	†	1,626	6.8 *	3.09	0.029	0.09
Students of Color	4,362	411	†	2,317	404	†	2,045	7.2 **	2.40	0.003	0.11
White	2,902	438	†	1,650	437	†	1,252	0.4	2.64	0.870	0.01
FRPL	3,435	409	†	1,892	402	†	1,543	7.0 ***	2.01	0.000	0.11
Not FRPL	2,712	448	†	1,622	446	†	1,090	1.9	3.15	0.542	0.02
ELs	1,091	393	†	440	387	†	651	5.2	4.09	0.200	0.09
Non-ELs	5,839	427	†	3,188	422	†	2,651	5.1	2.65	0.056	0.07
SPED	733	370	†	394	374	†	339	-4.3	2.24	0.055	-0.07
Non-SPED	6,250	430	†	3,287	425	†	2,963	4.5	2.49	0.070	0.06

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E10. Impacts on PSAT Math Scores, by Cohort and Subgroup, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	3,443	425	80	1,370	412	77	2,073	13.1 **	4.33	0.003	0.17
Subgroup											
Female	1,712	422	80	711	411	74	1,001	10.6 *	4.80	0.027	0.14
Male	1,731	427	79	659	412	78	1,072	15.5 *	7.50	0.039	0.20
Students of Color	1,652	413	77	546	394	75	1,106	19.4 **	6.67	0.004	0.26
White	1,789	438	79	823	428	74	966	10.0 **	3.51	0.004	0.13
FRPL	1,417	423	75	677	398	71	740	25.4 ***	7.61	0.001	0.35
Not FRPL	1,460	446	82	633	432	79	827	14.2 **	4.77	0.003	0.18
ELs	173	373	57	50	350	55	123	23.0	14.04	0.103	0.42
Non-ELs	2,928	425	81	1,198	412	73	1,730	13.5 **	4.67	0.004	0.18
SPED	321	363	64	98	368	71	223	-5.7	3.31	0.086	-0.08
Non-SPED	2,834	431	79	1,220	421	76	1,614	10.5 *	4.89	0.032	0.14
Cohort 2											
Full Sample	3,829	424	70	1,316	423	75	2,513	1.3	4.23	0.757	0.02
Subgroup											
Female	1,959	419	66	652	423	68	1,307	-3.8	4.76	0.423	-0.06
Male	1,870	429	74	664	423	81	1,206	6.0	4.42	0.177	0.08
Students of Color	2,710	414	60	885	407	68	1,825	6.8 *	3.43	0.049	0.10
White	1,113	430	77	430	474	80	683	-43.4 ***	8.58	0.000	-0.55
FRPL	2,018	410	63	681	403	64	1,337	6.3 *	2.93	0.032	0.10
Not FRPL	1,252	449	74	499	463	82	753	-13.9	8.89	0.119	-0.18
ELs	918	399	55	285	396	57	633	3.2	5.50	0.559	0.06
Non-ELs	2,911	431	72	1,031	432	77	1,880	-1.0	5.44	0.855	-0.01
SPED	412	374	58	161	378	56	251	-4.7	4.32	0.273	-0.08
Non-SPED	3,416	431	69	1,154	428	75	2,262	2.6	4.18	0.530	0.04

(continues)

Exhibit E10. Impacts on PSAT Math Scores, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	7,272	424	†	2,686	417	†	4,586	7.1 *	3.02	0.020	0.09
Subgroup											
Female	3,671	421	†	1,363	417	†	2,308	3.3	3.38	0.322	0.05
Male	3,601	428	†	1,323	420	†	2,278	8.4 *	3.81	0.027	0.11
Students of Color	4,362	414	†	1,431	404	†	2,931	9.4 **	3.05	0.002	0.14
White	2,902	437	†	1,253	434	†	1,649	2.4	3.25	0.466	0.03
FRPL	3,435	411	†	1,358	403	†	2,077	8.8 **	2.73	0.001	0.13
Not FRPL	2,712	447	†	1,132	439	†	1,580	7.9	4.20	0.059	0.10
ELs	1,091	396	†	335	390	†	756	5.8	5.12	0.253	0.10
Non-ELs	5,839	428	†	2,229	420	†	3,610	7.4 *	3.54	0.038	0.10
SPED	733	367	†	259	372	†	474	-5.3 *	2.63	0.042	-0.08
Non-SPED	6,250	431	†	2,374	425	†	3,876	5.9	3.18	0.061	0.08

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E11. Impacts on PSAT Total Scores, by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	3,443	854	157	2,193	832	146	1,250	21.6 *	8.83	0.014	0.14
Subgroup											
Female	1,712	857	156	1,101	846	144	611	11.0	7.92	0.166	0.07
Male	1,731	851	158	1,092	818	147	639	32.7 *	14.90	0.028	0.21
Students of Color	1,652	826	156	1,047	787	123	605	38.6 ***	11.26	0.001	0.27
White	1,789	881	154	1,144	872	150	645	9.1	5.57	0.103	0.06
FRPL	1,417	846	149	952	799	131	465	47.1 ***	12.45	0.000	0.33
Not FRPL	1,460	896	159	921	878	152	539	18.0	9.16	0.050	0.11
ELs	173	732	105	95	686	72	78	46.9 *	18.23	0.011	0.51
Non-ELs	2,928	858	155	1,756	837	146	1,172	21.2 *	9.16	0.020	0.14
SPED	321	735	134	186	736	107	135	-1.6	9.75	0.867	-0.01
Non-SPED	2,834	872	154	1,719	855	144	1,115	16.1	9.76	0.100	0.11
Cohort 2											
Full Sample	3,829	849	150	1,777	845	148	2,052	3.3	7.22	0.650	0.02
Subgroup											
Female	1,959	851	143	894	857	140	1,065	-5.2	6.99	0.458	-0.04
Male	1,870	845	157	883	833	154	987	11.7	8.14	0.149	0.08
Students of Color	2,710	816	137	1,270	806	120	1,440	9.8	6.49	0.130	0.08
White	1,113	915	154	506	952	161	607	-36.5 ***	7.13	0.000	-0.23
FRPL	2,018	810	124	940	800	121	1,078	10.3 *	4.45	0.021	0.08
Not FRPL	1,252	927	162	701	936	160	551	-9.0	10.89	0.407	-0.06
ELs	918	773	97	345	760	97	573	13.4	7.35	0.069	0.14
Non-ELs	2,911	872	154	1,432	873	152	1,479	-1.3	7.87	0.874	-0.01
SPED	412	736	110	208	740	99	204	-3.8	10.16	0.705	-0.04
Non-SPED	3,416	863	149	1,568	857	147	1,848	5.5	6.92	0.423	0.04

(continues)

Exhibit E11. Impacts on PSAT Total Scores, by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	7,272	851	†	3,970	840	†	3,302	10.6	5.59	0.057	0.07
Subgroup											
Female	3,671	854	†	1,995	852	†	1,676	1.9	5.24	0.718	0.01
Male	3,601	846	†	1,975	830	†	1,626	16.6 *	7.14	0.020	0.11
Students of Color	4,362	819	†	2,317	802	†	2,045	17.0 **	5.62	0.002	0.13
White	2,902	894	†	1,650	902	†	1,252	-8.2	4.39	0.061	-0.05
FRPL	3,435	814	†	1,892	799	†	1,543	14.4 ***	4.19	0.001	0.12
Not FRPL	2,712	909	†	1,622	902	†	1,090	6.8	7.01	0.334	0.04
ELs	1,091	767	†	440	749	†	651	18.1 **	6.82	0.008	0.19
Non-ELs	5,839	866	†	3,188	858	†	2,651	8.3	5.97	0.165	0.05
SPED	733	735	†	394	738	†	339	-2.7	7.04	0.701	-0.02
Non-SPED	6,250	866	†	3,287	857	†	2,963	9.1	5.65	0.109	0.06

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E12. Impacts on PSAT Total Scores, by Cohort and Subgroup, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	3,443	862	156	1,370	836	151	2,073	26.4 *	11.31	0.020	0.17
Subgroup											
Female	1,712	861	158	711	848	147	1,001	13.3	9.62	0.166	0.09
Male	1,731	863	154	659	824	153	1,072	39.8 *	19.48	0.041	0.26
Students of Color	1,652	844	153	546	796	142	1,106	48.4 ***	13.80	0.000	0.33
White	1,789	883	156	823	873	150	966	10.4	6.49	0.108	0.07
FRPL	1,417	862	151	677	802	140	740	59.5 ***	17.20	0.001	0.41
Not FRPL	1,460	901	158	633	880	156	827	21.2 *	10.48	0.044	0.13
ELs	173	755	102	50	694	87	123	61.5 *	28.88	0.035	0.67
Non-ELs	2,928	864	157	1,198	839	145	1,730	25.5 *	11.24	0.023	0.17
SPED	321	734	119	98	736	126	223	-1.8	10.44	0.863	-0.01
Non-SPED	2,834	876	155	1,220	857	149	1,614	19.1	11.81	0.105	0.13
Cohort 2											
Full Sample	3,829	850	142	1,316	845	152	2,513	4.5	9.51	0.640	0.03
Subgroup											
Female	1,959	850	135	652	857	145	1,307	-7.1	9.92	0.473	-0.05
Male	1,870	849	149	664	833	159	1,206	15.8	9.73	0.104	0.10
Students of Color	2,710	819	117	885	807	134	1,825	12.6	7.64	0.098	0.10
White	1,113	893	156	430	961	159	683	-67.6 ***	15.30	0.000	-0.43
FRPL	2,018	813	119	681	800	124	1,337	13.4 *	5.40	0.013	0.11
Not FRPL	1,252	921	152	499	937	166	753	-16.0	20.64	0.439	-0.10
ELs	918	776	98	285	760	96	633	16.1	9.36	0.085	0.17
Non-ELs	2,911	872	146	1,031	873	156	1,880	-1.9	11.69	0.874	-0.01
SPED	412	735	112	161	740	99	251	-5.4	13.61	0.692	-0.05
Non-SPED	3,416	865	140	1,154	857	152	2,262	7.6	8.84	0.392	0.05

(continues)

Exhibit E12. Impacts on PSAT Total Scores, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	7,272	855	†	2,686	841	†	4,586	13.5	7.28	0.063	0.09
Subgroup											
Female	3,671	856	†	1,363	852	†	2,308	3.4	6.90	0.621	0.02
Male	3,601	852	†	1,323	831	†	2,278	20.6 *	8.70	0.018	0.13
Students of Color	4,362	825	†	1,431	804	†	2,931	21.0 **	6.68	0.002	0.16
White	2,902	885	†	1,253	886	†	1,649	-1.5	5.97	0.806	-0.01
FRPL	3,435	818	†	1,358	800	†	2,077	17.5 ***	5.15	0.001	0.14
Not FRPL	2,712	905	†	1,132	892	†	1,580	13.6	9.34	0.147	0.09
ELs	1,091	774	†	335	753	†	756	20.4 *	8.90	0.022	0.21
Non-ELs	5,839	868	†	2,229	855	†	3,610	12.4	8.10	0.127	0.08
SPED	733	734	†	259	738	†	474	-3.1	8.28	0.705	-0.03
Non-SPED	6,250	869	†	2,374	857	†	3,876	11.7	7.08	0.098	0.08

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E13. Sensitivity Results for Impacts on PSAT Reading Scores, by Cohort and Subgroup

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
					Cohort 1						
Full Sample	4,122	427	86	2,872	424	84	1,250	2.4	3.64	0.505	0.03
Subgroup											
Female	2,051	436	85	1,440	432	84	611	3.8	4.34	0.383	0.04
Male	2,060	417	87	1,421	418	82	639	-0.3	3.17	0.912	0.00
Students of Color	2,069	413	85	1,464	405	72	605	8.4	5.92	0.158	0.10
White	1,934	444	87	1,289	443	87	645	0.5	3.06	0.868	0.01
FRPL	1,944	424	82	1,479	397	79	465	26.8***	7.74	0.001	0.33
Not FRPL	1,556	453	88	1,017	448	87	539	5.3	5.30	0.318	0.06
ELs	220	351	59	142	366	39	78	-14.9	18.73	0.428	-0.28
Non-ELs	3,541	430	85	2,369	426	84	1,172	3.9	3.74	0.293	0.05
SPED	395	368	72	260	359	58	135	8.8	7.75	0.255	0.13
Non-SPED	3,420	436	85	2,305	435	83	1,115	1.0	4.79	0.830	0.01
					Cohort 2						
Full Sample	3,881	425	90	1,777	422	88	2,104	2.3	4.84	0.633	0.03
Subgroup											
Female	1,988	431	88	894	433	87	1,094	-2.3	4.76	0.628	-0.03
Male	1,893	418	92	883	411	87	1,010	7.2	4.93	0.143	0.08
Students of Color	2,761	404	83	1,270	400	73	1,491	4.6	4.85	0.342	0.06
White	1,114	471	90	506	484	93	608	-13.1***	2.98	0.000	-0.14
FRPL	2,065	402	76	940	397	74	1,125	5.3	3.73	0.154	0.07
Not FRPL	1,270	472	96	714	473	94	556	-1.3	6.91	0.849	-0.01
ELs	922	375	59	345	364	56	577	10.7**	3.33	0.001	0.19
Non-ELs	2,959	440	91	1,432	441	89	1,527	-0.7	5.18	0.892	-0.01
SPED	417	361	72	208	361	62	209	-0.4	8.40	0.959	-0.01
Non-SPED	3,463	433	89	1,568	429	88	1,895	3.6	4.69	0.438	0.04

(continues)

Exhibit E13. Sensitivity Results for Impacts on PSAT Reading Scores, by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	8,003	426	†	4,649	424	†	3,354	2.4	2.91	0.412	0.03
Subgroup											
Female	4,039	434	†	2,334	433	†	1,705	1.0	3.21	0.750	0.01
Male	3,953	418	†	2,304	416	†	1,649	1.9	2.67	0.484	0.02
Students of Color	4,830	408	†	2,734	402	†	2,096	6.1	3.75	0.103	0.08
White	3,048	458	†	1,795	464	†	1,253	-6.5 **	2.13	0.002	-0.07
FRPL	4,009	406	†	2,419	397	†	1,590	9.4 **	3.36	0.005	0.12
Not FRPL	2,826	460	†	1,731	457	†	1,095	2.8	4.20	0.498	0.03
ELs	1,142	374	†	487	364	†	655	10.0 **	3.28	0.002	0.17
Non-ELs	6,500	434	†	3,801	431	†	2,699	2.3	3.03	0.439	0.03
SPED	812	365	†	468	360	†	344	4.6	5.70	0.422	0.07
Non-SPED	6,883	435	†	3,873	432	†	3,010	2.4	3.35	0.481	0.03

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. Sensitivity results include the three Cohort 1 schools and one Cohort 2 school that administered the PSAT-10 in place of the PSAT/NMSQT. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E14. Sensitivity Results for Impacts on PSAT Reading Scores, by Cohort and Subgroup, *Compliance Adjusted*

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
					Cohort 1						
Full Sample	4,122	428	89	1,407	425	83	2,715	3.2	4.81	0.506	0.04
Subgroup											
Female	2,051	438	89	732	433	81	1,319	4.9	5.66	0.384	0.06
Male	2,060	417	88	675	418	83	1,385	-0.5	4.19	0.912	-0.01
Students of Color	2,069	424	88	582	406	79	1,487	17.7	12.64	0.162	0.22
White	1,934	444	88	823	443	86	1,111	0.6	3.55	0.868	0.01
FRPL	1,944	439	87	677	405	78	1,267	33.9***	9.97	0.001	0.42
Not FRPL	1,556	455	89	633	448	87	923	6.3	6.27	0.318	0.07
ELs	220	344	63	50	360	49	170	-16.1	18.93	0.396	-0.30
Non-ELs	3,541	432	88	1,235	427	81	2,306	5.1	4.87	0.295	0.06
SPED	395	375	69	104	362	67	291	13.4	13.09	0.307	0.20
Non-SPED	3,420	437	88	1,251	436	82	2,169	1.3	5.98	0.830	0.02
					Cohort 2						
Full Sample	3,881	426	86	1,316	422	91	2,565	3.1	6.34	0.620	0.04
Subgroup											
Female	1,988	430	83	652	433	89	1,336	-3.2	6.71	0.637	-0.04
Male	1,893	421	88	664	411	90	1,229	9.7	5.98	0.105	0.11
Students of Color	2,761	406	73	885	400	80	1,876	5.9	5.79	0.307	0.08
White	1,114	463	91	430	487	92	684	-24.2**	7.88	0.002	-0.26
FRPL	2,065	404	74	681	397	76	1,384	6.9	4.52	0.126	0.09
Not FRPL	1,270	471	91	512	473	97	758	-2.3	12.36	0.851	-0.02
ELs	922	377	61	285	364	55	637	12.9**	4.32	0.003	0.23
Non-ELs	2,959	440	88	1,031	441	91	1,928	-1.0	7.69	0.892	-0.01
SPED	417	361	73	161	362	63	256	-0.6	11.45	0.958	-0.01
Non-SPED	3,463	434	85	1,154	429	90	2,309	5.0	5.95	0.405	0.06

(continues)

Exhibit E14. Sensitivity Results for Impacts on PSAT Reading Scores, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	8,003	427	†	2,723	424	†	5,280	3.2	3.83	0.407	0.04
Subgroup											
Female	4,039	435	†	1,384	433	†	2,655	1.6	4.33	0.717	0.02
Male	3,953	418	†	1,339	415	†	2,614	2.9	3.43	0.400	0.03
Students of Color	4,830	409	†	1,467	401	†	3,363	8.0	5.26	0.130	0.10
White	3,048	447	†	1,253	451	†	1,795	-3.6	3.23	0.268	-0.04
FRPL	4,009	410	†	1,358	399	†	2,651	11.5 **	4.12	0.005	0.15
Not FRPL	2,826	458	†	1,145	453	†	1,681	4.5	5.59	0.421	0.05
ELs	1,142	375	†	335	364	†	807	11.5 **	4.21	0.006	0.20
Non-ELs	6,500	435	†	2,266	431	†	4,234	3.3	4.11	0.416	0.04
SPED	812	367	†	265	362	†	547	5.5	8.62	0.526	0.08
Non-SPED	6,883	436	†	2,405	433	†	4,478	3.1	4.22	0.458	0.04

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. Sensitivity results include the three Cohort 1 schools and one Cohort 2 school that administered the PSAT-10 in place of the PSAT/NMSQT. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E15. Sensitivity Results for Impacts on PSAT Math Scores, by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	4,122	416	78	2,872	409	74	1,250	6.8	3.66	0.062	0.09
Subgroup											
Female	2,051	417	76	1,440	406	72	611	11.1 *	5.50	0.044	0.15
Male	2,060	415	80	1,421	413	76	639	2.0	4.03	0.625	0.03
Students of Color	2,069	401	77	1,464	396	65	605	4.4	6.00	0.459	0.06
White	1,934	433	77	1,289	425	75	645	8.6 **	3.18	0.007	0.11
FRPL	1,944	411	73	1,479	391	65	465	20.1 ***	5.65	0.000	0.28
Not FRPL	1,556	441	81	1,017	429	76	539	11.8 **	4.44	0.008	0.15
ELs	220	347	55	142	380	55	78	-32.9	25.92	0.206	-0.60
Non-ELs	3,541	417	76	2,369	410	74	1,172	7.6	4.04	0.061	0.10
SPED	395	362	71	260	362	64	135	-0.1	3.93	0.974	0.00
Non-SPED	3,420	425	76	2,305	418	73	1,115	6.7	5.23	0.198	0.09
Cohort 2											
Full Sample	3,881	424	73	1,777	423	72	2,104	0.9	3.17	0.779	0.01
Subgroup											
Female	1,988	420	68	894	423	66	1,094	-2.7	3.45	0.426	-0.04
Male	1,893	427	78	883	423	79	1,010	4.3	3.67	0.239	0.06
Students of Color	2,761	413	68	1,270	408	63	1,491	5.2 *	2.63	0.049	0.08
White	1,114	444	77	506	468	80	608	-23.4 ***	5.20	0.000	-0.30
FRPL	2,065	409	65	940	404	63	1,125	4.7 *	2.12	0.027	0.07
Not FRPL	1,270	453	79	714	461	79	556	-7.8	4.42	0.078	-0.10
ELs	922	398	54	345	396	57	577	2.7	4.48	0.543	0.05
Non-ELs	2,959	431	76	1,432	432	75	1,527	-0.8	3.68	0.819	-0.01
SPED	417	375	57	208	379	57	209	-3.4	3.19	0.286	-0.06
Non-SPED	3,463	430	73	1,568	428	72	1,895	1.8	3.16	0.561	0.03

(continues)

Exhibit E15. Sensitivity Results for Impacts on PSAT Math Scores, by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	8,003	420	†	4,649	417	†	3,354	3.4	2.39	0.152	0.05
Subgroup											
Female	4,039	419	†	2,334	418	†	1,705	1.1	2.92	0.694	0.02
Male	3,953	422	†	2,304	418	†	1,649	3.3	2.71	0.230	0.04
Students of Color	4,830	411	†	2,734	406	†	2,096	5.1 *	2.41	0.035	0.08
White	3,048	436	†	1,795	436	†	1,253	-0.1	2.71	0.977	0.00
FRPL	4,009	409	†	2,419	402	†	1,590	6.6 ***	1.98	0.001	0.10
Not FRPL	2,826	447	†	1,731	445	†	1,095	2.0	3.13	0.533	0.02
ELs	1,142	397	†	487	395	†	655	1.7	4.42	0.702	0.03
Non-ELs	6,500	425	†	3,801	422	†	2,699	3.0	2.72	0.273	0.04
SPED	812	370	†	468	372	†	344	-2.1	2.48	0.395	-0.03
Non-SPED	6,883	428	†	3,873	425	†	3,010	3.1	2.70	0.245	0.04

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. Sensitivity results include the three Cohort 1 schools and one Cohort 2 school that administered the PSAT-10 in place of the PSAT/NMSQT. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E16. Sensitivity Results for Impacts on PSAT Math Scores, by Cohort and Subgroup, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	4,122	420	79	1,407	411	75	2,715	9.0 *	4.53	0.047	0.12
Subgroup											
Female	2,051	423	80	732	408	71	1,319	14.4 *	7.02	0.040	0.19
Male	2,060	416	79	675	414	78	1,385	2.6	5.16	0.613	0.03
Students of Color	2,069	406	77	582	397	74	1,487	9.4	11.29	0.405	0.13
White	1,934	436	79	823	426	73	1,111	10.0 **	3.71	0.007	0.13
FRPL	1,944	422	75	677	397	69	1,267	25.3 ***	7.67	0.001	0.36
Not FRPL	1,556	445	82	633	431	78	923	13.9 **	4.72	0.003	0.18
ELs	220	331	57	50	367	55	170	-35.6	26.87	0.187	-0.64
Non-ELs	3,541	421	81	1,235	411	72	2,306	9.8 *	4.90	0.045	0.13
SPED	395	361	63	104	362	70	291	-0.2	5.80	0.973	0.00
Non-SPED	3,420	428	79	1,251	419	73	2,169	8.4	6.35	0.185	0.11
Cohort 2											
Full Sample	3,881	424	70	1,316	423	74	2,565	1.2	4.24	0.78	0.02
Subgroup											
Female	1,988	419	66	652	423	68	1,336	-3.8	4.74	0.43	-0.06
Male	1,893	429	74	664	423	81	1,229	5.8	4.45	0.19	0.07
Students of Color	2,761	414	60	885	408	68	1,876	6.7 *	3.35	0.05	0.10
White	1,114	430	77	430	474	80	684	-43.4 ***	8.57	0.00	-0.55
FRPL	2,065	410	63	681	404	64	1,384	6.1 *	2.82	0.03	0.10
Not FRPL	1,270	449	74	512	462	82	758	-13.8	8.91	0.12	-0.18
ELs	922	399	55	285	396	57	637	3.3	5.47	0.55	0.06
Non-ELs	2,959	431	72	1,031	432	77	1,928	-1.2	5.45	0.82	-0.02
SPED	417	374	58	161	379	57	256	-4.8	4.31	0.27	-0.08
Non-SPED	3,463	431	69	1,154	428	74	2,309	2.5	4.19	0.55	0.03

(continues)

Exhibit E16. Sensitivity Results for Impacts on PSAT Math Scores, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	8,003	422	†	2,723	417	†	5,280	4.9	3.09	0.12	0.07
Subgroup											
Female	4,039	420	†	1,384	418	†	2,655	1.9	3.93	0.62	0.03
Male	3,953	423	†	1,339	419	†	2,614	4.5	3.37	0.19	0.06
Students of Color	4,830	414	†	1,467	407	†	3,363	6.9 *	3.21	0.03	0.10
White	3,048	435	†	1,253	434	†	1,795	1.6	3.40	0.64	0.02
FRPL	4,009	412	†	1,358	403	†	2,651	8.4 **	2.65	0.00	0.13
Not FRPL	2,826	446	†	1,145	438	†	1,681	7.8	4.17	0.06	0.10
ELs	1,142	396	†	335	395	†	807	1.7	5.36	0.75	0.03
Non-Els	6,500	425	†	2,266	421	†	4,234	4.9	3.65	0.18	0.07
SPED	812	370	†	265	373	†	547	-3.1	3.46	0.36	-0.05
Non-SPED	6,883	430	†	2,405	425	†	4,478	4.3	3.50	0.22	0.06

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. Sensitivity results include the three Cohort 1 schools and one Cohort 2 school that administered the PSAT-10 in place of the PSAT/NMSQT. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E17. Sensitivity Results for Impacts on PSAT Total Scores, by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	4,122	842	152	2,872	833	146	1,250	9.2	6.80	0.177	0.06
Subgroup											
Female	2,051	852	148	1,440	838	144	611	14.7	9.16	0.109	0.10
Male	2,060	833	155	1,421	831	147	639	1.6	5.97	0.786	0.01
Students of Color	2,069	814	150	1,464	801	123	605	12.8	10.74	0.234	0.09
White	1,934	877	152	1,289	868	150	645	9.0	5.85	0.123	0.06
FRPL	1,944	834	143	1,479	787	131	465	46.8***	13.30	0.000	0.33
Not FRPL	1,556	894	157	1,017	877	152	539	17.1	8.95	0.057	0.11
ELs	220	698	99	142	746	72	78	-47.8	43.72	0.276	-0.53
Non-ELs	3,541	848	149	2,369	836	146	1,172	11.4	7.34	0.119	0.08
SPED	395	730	129	260	721	107	135	8.7	10.54	0.409	0.07
Non-SPED	3,420	861	149	2,305	853	144	1,115	7.7	9.66	0.428	0.05
Cohort 2											
Full Sample	3,881	849	150	1,777	845	147	2,104	3.2	7.26	0.659	0.02
Subgroup											
Female	1,988	851	143	894	856	139	1,094	-5.1	7.02	0.472	-0.04
Male	1,893	845	157	883	834	153	1,010	11.5	8.16	0.157	0.07
Students of Color	2,761	817	137	1,270	807	120	1,491	9.8	6.48	0.131	0.08
White	1,114	915	154	506	951	161	608	-36.5***	7.12	0.000	-0.23
FRPL	2,065	811	124	940	801	121	1,125	10.0*	4.46	0.025	0.08
Not FRPL	1,270	925	162	714	934	160	556	-9.1	10.89	0.402	-0.06
ELs	922	773	97	345	760	97	577	13.5	7.31	0.066	0.14
Non-ELs	2,959	872	154	1,432	873	151	1,527	-1.5	7.91	0.845	-0.01
SPED	417	737	110	208	740	99	209	-3.8	10.18	0.707	-0.04
Non-SPED	3,463	863	149	1,568	857	146	1,895	5.5	6.97	0.433	0.04

(continues)

Exhibit E17. Sensitivity Results for Impacts on PSAT Total Scores, by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	8,003	845	†	4,649	839	†	3,354	6.4	4.96	0.198	0.04
Subgroup											
Female	4,039	852	†	2,334	850	†	1,705	2.2	5.57	0.687	0.02
Male	3,953	837	†	2,304	832	†	1,649	5.1	4.82	0.292	0.03
Students of Color	4,830	816	†	2,734	805	†	2,096	10.6	5.55	0.056	0.08
White	3,048	892	†	1,795	902	†	1,253	-9.3 *	4.52	0.039	-0.06
FRPL	4,009	813	†	2,419	800	†	1,590	13.7 **	4.23	0.001	0.11
Not FRPL	2,826	906	†	1,731	900	†	1,095	6.5	6.91	0.346	0.04
ELs	1,142	771	†	487	759	†	655	11.8	7.21	0.102	0.12
Non-ELs	6,500	859	†	3,801	853	†	2,699	5.4	5.38	0.313	0.04
SPED	812	733	†	468	731	†	344	2.2	7.32	0.762	0.02
Non-SPED	6,883	862	†	3,873	856	†	3,010	6.2	5.65	0.271	0.04

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. Sensitivity results include the three Cohort 1 schools and one Cohort 2 school that administered the PSAT-10 in place of the PSAT/NMSQT. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E18. Sensitivity Results for Impacts on PSAT Total Scores, by Cohort and Subgroup, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	4,122	848	156	1,407	836	145	2,715	12.1	8.79	0.168	0.08
Subgroup											
Female	2,051	860	158	732	841	140	1,319	19.2	11.86	0.107	0.13
Male	2,060	833	154	675	831	150	1,385	2.1	7.79	0.783	0.01
Students of Color	2,069	830	152	582	802	139	1,487	27.1	20.64	0.189	0.19
White	1,934	880	156	823	870	147	1,111	10.5	6.90	0.129	0.07
FRPL	1,944	862	151	677	802	134	1,267	59.1***	17.53	0.001	0.42
Not FRPL	1,556	900	158	633	880	153	923	20.2*	10.26	0.049	0.13
ELs	220	675	102	50	727	87	170	-51.6	44.65	0.249	-0.57
Non-ELs	3,541	853	157	1,235	839	141	2,306	14.8	9.29	0.110	0.10
SPED	395	736	120	104	723	123	291	13.2	17.43	0.449	0.11
Non-SPED	3,420	865	155	1,251	855	143	2,169	9.6	11.97	0.424	0.07
Cohort 2											
Full Sample	3,881	850	142	1,316	845	152	2,565	4.3	9.55	0.649	0.03
Subgroup											
Female	1,988	850	135	652	856	144	1,336	-6.9	9.94	0.486	-0.05
Male	1,893	849	149	664	834	159	1,229	15.5	9.76	0.112	0.10
Students of Color	2,761	820	117	885	808	134	1,876	12.6	7.61	0.099	0.10
White	1,114	893	156	430	961	159	684	-67.6***	15.27	0.000	-0.43
FRPL	2,065	814	119	681	801	124	1,384	13.0*	5.38	0.015	0.11
Not FRPL	1,270	919	153	512	936	167	758	-16.1	20.66	0.435	-0.10
ELs	922	776	98	285	760	96	637	16.2	9.30	0.082	0.17
Non-ELs	2,959	871	146	1,031	873	155	1,928	-2.3	11.77	0.846	-0.01
SPED	417	735	112	161	741	99	256	-5.4	13.62	0.694	-0.05
Non-SPED	3,463	865	140	1,154	857	151	2,309	7.5	8.90	0.402	0.05

(continues)

Exhibit E18. Sensitivity Results for Impacts on PSAT Total Scores, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	8,003	849	†	2,723	840	†	5,280	8.5	6.46	0.186	0.06
Subgroup											
Female	4,039	854	†	1,384	850	†	2,655	3.8	7.62	0.615	0.03
Male	3,953	840	†	1,339	832	†	2,614	7.4	6.09	0.227	0.05
Students of Color	4,830	821	†	1,467	807	†	3,363	14.3 *	7.14	0.045	0.11
White	3,048	882	†	1,253	885	†	1,795	-2.8	6.29	0.659	-0.02
FRPL	4,009	818	†	1,358	801	†	2,651	17.0 ***	5.14	0.001	0.14
Not FRPL	2,826	904	†	1,145	891	†	1,681	13.0	9.19	0.157	0.08
ELs	1,142	772	†	335	758	†	807	13.4	9.10	0.142	0.14
Non-ELs	6,500	860	†	2,266	852	†	4,234	8.3	7.29	0.257	0.06
SPED	812	736	†	265	734	†	547	1.7	10.73	0.876	0.02
Non-SPED	6,883	865	†	2,405	857	†	4,478	8.2	7.14	0.250	0.06

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. Sensitivity results include the three Cohort 1 schools and one Cohort 2 school that administered the PSAT-10 in place of the PSAT/NMSQT. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations for these estimates.

Source: AIR calculations from school-provided administrative data.

Exhibit E19. Impacts on Chronic Absenteeism (i.e., Percentage of Students Chronically Absent), by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	6,303	12.6	38.9	4,246	18.0	37.6	2,057	-5.42	23.49	0.072	-0.26
Subgroup											
Female	2,951	11.9	38.2	1,973	16.2	38.1	978	-4.35	27.08	0.180	-0.22
Male	3,168	12.3	38.7	2,089	18.5	37.2	1,079	-6.26 *	23.90	0.042	-0.29
Students of Color	3,416	10.2	39.9	2,213	21.2	36.5	1,203	-11.03 *	42.35	0.041	-0.52
White	2,546	13.5	37.0	1,692	15.4	39.1	854	-1.90	18.60	0.408	-0.09
FRPL	2,742	13.9	36.0	2,071	19.3	43.4	671	-5.43	37.87	0.296	-0.24
Not FRPL	1,906	8.4	28.2	1,244	9.2	34.1	662	-0.82	16.60	0.539	-0.06
ELs	423	2.7	36.6	252	22.5	31.5	171	-19.81 ***	35.54	0.000	-1.43
Non-ELs	5,220	13.4	39.2	3,339	16.7	37.9	1,881	-3.33	24.22	0.281	-0.16
SPED	898	23.1	45.7	550	25.2	42.5	348	-2.13	24.84	0.640	-0.07
Non-SPED	4,697	9.7	34.7	2,993	13.0	36.2	1,704	-3.28	31.22	0.294	-0.20
Cohort 2											
Full Sample	7,533	24.1	44.7	3,568	24.6	44.6	3,965	-0.43	14.02	0.866	-0.01
Subgroup											
Female	3,602	22.3	44.1	1,701	24.1	44.4	1,901	-1.80	15.34	0.511	-0.06
Male	3,915	25.1	45.1	1,860	24.6	44.6	2,055	0.53	15.40	0.854	0.02
Students of Color	5,934	27.9	46.1	2,921	27.8	45.6	3,013	0.13	16.84	0.969	0.00
White	1,565	13.5	33.7	636	11.4	40.0	929	2.10	23.66	0.412	0.12
FRPL	4,636	30.7	47.6	2,323	30.3	45.8	2,313	0.37	16.54	0.916	0.01
Not FRPL	1,869	10.4	34.6	1,009	12.0	36.7	860	-1.61	14.17	0.253	-0.10
ELs	1,938	25.5	44.6	809	26.8	45.2	1,129	-1.29	11.46	0.560	-0.04
Non-ELs	5,578	22.9	44.7	2,752	22.9	44.2	2,826	0.02	16.72	0.994	0.00
SPED	1,171	39.3	49.4	534	39.2	48.8	637	0.08	16.51	0.985	0.00
Non-SPED	6,344	21.5	43.3	3,026	21.9	43.3	3,318	-0.47	15.11	0.856	-0.02

(continues)

Exhibit E19. Impacts on Chronic Absenteeism (i.e., Percentage of Students Chronically Absent), by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 3											
Full Sample	7,519	13.2	39.7	4,236	19.0	41.5	3,283	-5.89 *	20.10	0.028	-0.27
Subgroup											
Female	3,621	14.0	39.9	2,064	17.3	40.9	1,557	-3.39	20.91	0.218	-0.16
Male	3,885	12.0	39.4	2,170	20.2	41.7	1,715	-8.16 **	20.49	0.003	-0.37
Students of Color	4,784	13.8	39.1	2,610	21.0	43.8	2,174	-7.28 *	20.09	0.011	-0.31
White	2,716	12.6	40.5	1,623	11.0	34.6	1,093	1.65	10.27	0.121	0.10
FRPL	3,542	15.7	38.0	2,098	21.5	43.9	1,444	-5.81 *	18.29	0.035	-0.23
Not FRPL	2,150	8.3	30.2	821	8.1	31.7	1,329	0.25	13.51	0.807	0.02
ELs	923	19.6	40.8	579	20.4	44.5	344	-0.82	25.17	0.838	-0.03
Non-ELs	6,508	11.6	39.5	3,652	17.6	40.1	2,856	-6.02 **	18.67	0.009	-0.30
SPED	1,060	28.4	46.7	650	30.6	46.4	410	-2.19	19.30	0.586	-0.06
Non-SPED	6,371	10.7	37.8	3,581	16.5	39.5	2,790	-5.78 *	21.13	0.019	-0.30
Cohorts 1, 2, and 3											
Full Sample	21,355	19.0	†	12,050	21.8	†	9,305	-2.84 *	10.33	0.041	-0.11
Subgroup											
Female	10,174	18.1	†	5,738	20.8	†	4,436	-2.70	11.25	0.089	-0.11
Male	10,968	18.7	†	6,119	22.0	†	4,849	-3.37 *	10.94	0.016	-0.13
Students of Color	14,134	21.1	†	7,744	24.7	†	6,390	-3.62 *	12.34	0.033	-0.13
White	6,827	12.9	†	3,951	11.9	†	2,876	0.98	8.40	0.235	0.06
FRPL	10,920	23.0	†	6,492	25.7	†	4,428	-2.70	11.67	0.111	-0.09
Not FRPL	5,925	9.1	†	3,074	9.8	†	2,851	-0.68	8.42	0.401	-0.05
ELs	3,284	22.7	†	1,640	25.4	†	1,644	-2.69 *	10.01	0.014	-0.09
Non-ELs	17,306	17.0	†	9,743	19.8	†	7,563	-2.80 *	11.08	0.041	-0.12
SPED	3,129	32.3	†	1,734	33.5	†	1,395	-1.13	11.20	0.607	-0.03
Non-SPED	17,412	16.7	†	9,600	19.1	†	7,812	-2.40	11.44	0.073	-0.10

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations.

Source: AIR calculations from school-provided administrative data.

Exhibit E20. Impacts on Chronic Absenteeism (i.e., Percentage of Students Chronically Absent), by Cohort and Subgroup, *Compliance Adjusted*

		Treatment			Control						Effect
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Size
					Cohort 1						
Full Sample	6,303	8.9	39.7	1,948	17.3	37.9	4,355	-8.39	42.30	0.173	-0.46
Subgroup											
Female	2,951	9.1	39.7	983	15.7	37.3	1,968	-6.66	46.73	0.280	-0.38
Male	3,168	8.1	39.6	964	17.9	37.5	2,204	-9.76	44.41	0.141	-0.55
Students of Color	3,416	1.0	42.4	875	27.9	37.3	2,541	-26.87	176.55	0.293	-2.18
White	2,546	12.9	37.1	1,067	15.1	38.1	1,479	-2.26	22.78	0.418	-0.11
FRPL	2,742	11.0	40.7	953	17.7	36.7	1,789	-6.73	53.30	0.322	-0.34
Not FRPL	1,906	8.1	29.0	768	9.0	31.4	1,138	-0.96	20.04	0.630	-0.07
ELs	423	0.0	34.7	80	39.2	34.7	343	-39.22	175.66	0.198	-7.07
Non-ELs	5,220	11.4	40.4	1,710	16.3	37.9	3,510	-4.82	37.74	0.439	-0.25
SPED	898	20.9	49.0	217	24.9	42.3	681	-4.04	48.97	0.701	-0.14
Non-SPED	4,697	8.0	37.2	1,638	12.7	34.2	3,059	-4.63	48.25	0.439	-0.31
					Cohort 2						
Full Sample	7,533	23.9	43.2	1,988	24.5	45.1	5,545	-0.64	20.87	0.893	-0.02
Subgroup											
Female	3,602	21.3	43.0	938	24.0	44.7	2,664	-2.67	23.24	0.695	-0.09
Male	3,915	25.3	43.3	1,050	24.6	45.4	2,865	0.78	22.45	0.612	0.03
Students of Color	5,934	28.0	45.4	1,453	27.8	46.0	4,481	0.19	24.45	0.821	0.01
White	1,565	14.4	33.2	531	11.0	39.7	1,034	3.37	37.37	0.323	0.19
FRPL	4,636	30.9	47.3	1,087	30.3	46.6	3,549	0.55	24.40	0.881	0.02
Not FRPL	1,869	9.3	33.9	672	12.1	36.5	1,197	-2.77	25.52	0.856	-0.18
ELs	1,938	24.8	44.5	572	26.5	45.2	1,366	-1.66	14.94	0.567	-0.05
Non-ELs	5,578	23.0	42.6	1,416	23.0	45.0	4,162	0.04	27.19	0.776	0.00
SPED	1,171	39.4	47.4	299	39.3	49.5	872	0.12	26.16	0.903	0.00
Non-SPED	6,344	21.1	42.2	1,688	21.8	43.7	4,656	-0.68	22.22	0.899	-0.02

(continues)

Exhibit E20. Impacts on Chronic Absenteeism, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 3											
Full Sample	7,519	12.7	40.5	3,942	19.2	40.5	3,577	-6.45 *	22.21	0.043	-0.30
Subgroup											
Female	3,621	13.7	40.6	1,929	17.4	40.0	1,692	-3.70	22.94	0.202	-0.17
Male	3,885	11.4	40.3	2,011	20.4	40.7	1,874	-8.94 **	22.79	0.009	-0.41
Students of Color	4,784	13.2	39.8	2,436	21.2	43.0	2,348	-7.98 *	22.27	0.011	-0.34
White	2,716	12.8	41.5	1,503	10.9	33.6	1,213	1.90 **	11.70	0.003	0.11
FRPL	3,542	15.6	38.1	2,046	21.5	43.7	1,496	-5.89 *	18.62	0.030	-0.24
Not FRPL	2,150	8.4	31.2	732	8.1	31.1	1,418	0.29	15.50	0.753	0.02
ELs	923	19.6	41.1	560	20.4	44.0	363	-0.83	25.32	0.935	-0.03
Non-ELs	6,508	11.1	40.3	3,377	17.8	39.1	3,131	-6.64 *	20.87	0.020	-0.33
SPED	1,060	28.3	47.0	626	30.5	46.1	434	-2.24	19.77	0.695	-0.07
Non-SPED	6,371	10.3	38.6	3,311	16.7	38.5	3,060	-6.40 *	23.64	0.030	-0.34
Cohorts 1, 2, and 3											
Full Sample	21,355	17.5	†	7,878	21.5	†	13,477	-3.94 *	14.31	0.033	-0.16
Subgroup											
Female	10,174	16.6	†	3,850	20.1	†	6,324	-3.57	15.41	0.088	-0.15
Male	10,968	17.3	†	4,025	22.0	†	6,943	-4.67 *	15.05	0.011	-0.18
Students of Color	14,134	19.8	†	4,764	24.2	†	9,370	-4.47 *	16.39	0.041	-0.16
White	6,827	13.0	†	3,101	11.8	†	3,726	1.20	10.02	0.236	0.07
FRPL	10,920	20.5	†	4,086	24.2	†	6,834	-3.75	14.26	0.066	-0.13
Not FRPL	5,925	8.5	†	2,172	9.1	†	3,753	-0.67	11.05	0.509	-0.05
ELs	3,284	23.4	†	1,212	25.0	†	2,072	-1.65	12.84	0.276	-0.05
Non-ELs	17,306	14.9	†	6,503	19.1	†	10,803	-4.27 *	15.16	0.020	-0.19
SPED	3,129	31.3	†	1,142	32.9	†	1,987	-1.63	15.01	0.585	-0.05
Non-SPED	17,412	15.3	†	6,637	18.7	†	10,775	-3.49 *	15.35	0.047	-0.15

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations.

Source: AIR calculations from school-provided administrative data.

Exhibit E21. Impacts on Suspensions (i.e., Percentage of Students Suspended), by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	5,841	7.6	32.6	3,782	11.6	32.4	2,059	-4.04 *	21.90	0.031	-0.29
Subgroup											
Female	2,727	6.2	28.4	1,747	7.9	27.5	980	-1.68	35.72	0.472	-0.16
Male	2,930	7.9	34.0	1,851	14.6	35.9	1,079	-6.74 ***	19.47	0.000	-0.42
Students of Color	2,996	6.4	32.6	1,792	11.5	30.9	1,204	-5.08	42.02	0.128	-0.39
White	2,508	6.3	29.2	1,653	9.7	34.3	855	-3.43 **	14.43	0.001	-0.29
FRPL	2,742	11.2	34.9	2,071	19.2	39.6	671	-8.04 *	32.46	0.050	-0.39
Not FRPL	1,908	3.8	22.3	1,244	7.8	31.3	664	-4.07 ***	21.10	0.000	-0.47
ELs	324	0.6	33.1	153	2.3	32.9	171	-1.70 **	49.56	0.005	-0.84
Non-ELs	4,857	3.0	31.7	2,974	4.9	32.4	1,883	-1.87 *	23.06	0.029	-0.30
SPED	899	15.0	39.9	550	22.2	41.5	349	-7.21	28.17	0.088	-0.29
Non-SPED	4,698	5.3	28.7	2,993	8.3	29.8	1,705	-3.01 *	22.21	0.029	-0.29
Cohort 2											
Full Sample	5,671	9.3	29.0	2,328	8.4	32.8	3,343	0.90	19.65	0.568	0.07
Subgroup											
Female	2,738	6.5	27.6	1,125	6.1	29.3	1,613	0.40	24.60	0.782	0.04
Male	2,917	12.5	30.4	1,196	11.0	35.7	1,721	1.48	20.38	0.481	0.09
Students of Color	4,077	11.9	30.4	1,696	11.2	35.0	2,381	0.77	22.22	0.736	0.05
White	1,560	3.9	24.9	621	4.1	26.1	939	-0.20	27.30	0.845	-0.03
FRPL	2,990	13.4	34.1	1,189	11.7	34.8	1,801	1.74	19.09	0.406	0.10
Not FRPL	1,653	3.7	21.5	903	3.3	20.8	750	0.39	17.21	0.505	0.07
ELs	1,446	10.8	27.6	542	11.4	31.1	904	-0.56	18.18	0.753	-0.03
Non-ELs	4,208	8.8	29.5	1,779	7.9	33.5	2,429	0.89	21.94	0.597	0.07
SPED	865	17.3	39.4	345	12.7	38.4	520	4.59 *	17.97	0.043	0.22
Non-SPED	4,788	7.8	26.5	1,975	7.6	31.7	2,813	0.21	22.87	0.899	0.02

(continues)

Exhibit E21. Impacts on Suspensions (i.e., Percentage of Students Suspended), by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 3											
Full Sample	7,392	4.2	24.2	4,198	3.3	25.0	3,194	0.98	23.17	0.241	0.16
Subgroup											
Female	3,574	2.7	20.4	2,053	1.9	23.4	1,521	0.73	23.64	0.163	0.20
Male	3,805	5.8	27.2	2,143	4.7	26.5	1,662	1.19	27.40	0.380	0.15
Students of Color	4,630	4.4	24.3	2,556	4.3	27.4	2,074	0.16	22.52	0.866	0.02
White	2,743	3.9	24.1	1,639	1.9	19.8	1,104	1.94 ***	18.11	0.000	0.43
FRPL	3,374	5.0	27.1	2,037	3.7	27.9	1,337	1.31	20.26	0.122	0.19
Not FRPL	2,154	1.9	16.2	819	1.7	21.5	1,335	0.20	35.31	0.747	0.07
ELs	892	2.2	22.7	551	5.4	27.9	341	-3.17 *	37.27	0.014	-0.56
Non-ELs	6,411	4.3	24.4	3,642	3.0	24.7	2,769	1.29	20.08	0.062	0.23
SPED	1,017	13.8	35.4	634	8.4	34.6	383	5.38 *	24.16	0.022	0.34
Non-SPED	6,287	3.1	21.3	3,559	2.7	23.2	2,728	0.34	23.65	0.609	0.07
Cohorts 1, 2, and 3											
Full Sample	18,904	7.3	†	10,308	8.0	†	8,596	-0.65	12.37	0.818	-0.06
Subgroup											
Female	9,039	4.8	†	4,925	4.7	†	4,114	0.16	15.38	0.441	0.02
Male	9,652	9.2	†	5,190	11.2	†	4,462	-1.98	12.52	0.144	-0.13
Students of Color	11,703	8.0	†	6,044	8.2	†	5,659	-0.22	14.80	0.841	-0.02
White	6,811	5.2	†	3,913	6.3	†	2,898	-1.18	10.43	0.861	-0.14
FRPL	9,106	9.7	†	5,297	9.7	†	3,809	0.05	12.77	0.448	0.00
Not FRPL	5,715	3.5	†	2,966	4.7	†	2,749	-1.19	12.48	0.115	-0.19
ELs	2,662	8.3	†	1,246	9.5	†	1,416	-1.13 *	15.52	0.030	-0.09
Non-ELs	15,476	5.4	†	8,395	5.1	†	7,081	0.24	12.46	0.780	0.03
SPED	2,781	15.8	†	1,529	13.5	†	1,252	2.36	12.84	0.060	0.12
Non-SPED	15,773	5.4	†	8,527	6.3	†	7,246	-0.89	13.21	0.348	-0.10

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations.

Source: AIR calculations from school-provided administrative data.

Exhibit E22. Impacts on Suspensions (i.e., Percentage of Students Suspended), by Cohort and Subgroup, *Compliance Adjusted*

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	5,841	5.2	28.8	1,864	10.5	34.0	3,977	-5.30 *	35.21	0.031	-0.46
Subgroup											
Female	2,727	4.9	26.0	934	7.2	29.1	1,793	-2.27	56.14	0.537	-0.24
Male	2,930	4.8	31.1	929	13.5	36.1	2,001	-8.63 **	31.36	0.004	-0.68
Students of Color	2,996	2.1	31.2	798	12.0	32.2	2,198	-9.89 *	120.61	0.049	-1.11
White	2,508	5.4	26.7	1,061	9.1	33.7	1,447	-3.70 ***	17.11	0.000	-0.34
FRPL	2,742	7.6	32.0	953	16.7	38.1	1,789	-9.15 ***	45.70	0.000	-0.54
Not FRPL	1,908	2.9	22.8	768	7.0	27.8	1,140	-4.11 ***	25.47	0.001	-0.57
ELs	324	0.0	27.9	60	6.8	34.0	264	-6.78	242.40	0.103	-4.09
Non-ELs	4,857	5.0	28.7	1,646	10.3	33.4	3,211	-5.32 *	35.84	0.030	-0.47
SPED	899	9.4	42.5	217	21.1	39.9	682	-11.70	55.48	0.141	-0.57
Non-SPED	4,698	3.9	25.9	1,638	7.9	30.6	3,060	-4.01 *	34.32	0.015	-0.45
Cohort 2											
Full Sample	5,671	8.8	27.8	1,606	7.6	32.6	4,065	1.22	28.48	0.944	0.10
Subgroup											
Female	2,738	6.6	24.5	764	6.0	30.0	1,974	0.59	36.19	0.925	0.06
Male	2,917	11.3	30.3	842	9.4	34.9	2,075	1.88	28.98	0.950	0.12
Students of Color	4,077	10.5	28.6	1,069	9.6	34.6	3,008	0.96	31.32	0.961	0.06
White	1,560	3.9	25.8	533	4.3	25.5	1,027	-0.33	42.80	0.373	-0.05
FRPL	2,990	13.6	32.7	734	11.1	35.1	2,256	2.51	28.06	0.645	0.14
Not FRPL	1,653	3.0	22.4	643	2.5	20.4	1,010	0.54	30.46	0.443	0.12
ELs	1,446	7.7	27.6	422	8.3	30.7	1,024	-0.52	22.67	0.395	-0.04
Non-ELs	4,208	8.9	27.8	1,184	7.5	33.3	3,024	1.38	34.85	0.821	0.11
SPED	865	22.2	35.8	213	13.8	39.6	652	8.46	28.75	0.185	0.35
Non-SPED	4,788	6.9	26.2	1,392	6.6	31.0	3,396	0.26	32.63	0.671	0.03

(continues)

Exhibit E22. Impacts on Suspensions, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 3											
Full Sample	7,392	4.3	24.6	3,904	3.2	24.5	3,488	1.09	25.82	0.739	0.18
Subgroup											
Female	3,574	2.7	20.9	1,918	1.9	22.6	1,656	0.81	26.11	0.468	0.22
Male	3,805	6.0	27.5	1,984	4.6	26.2	1,821	1.34	30.79	0.994	0.16
Students of Color	4,630	4.3	24.6	2,382	4.2	26.9	2,248	0.17	25.33	0.295	0.03
White	2,743	4.2	24.6	1,519	1.9	19.4	1,224	2.27 *	20.60	0.016	0.49
FRPL	3,374	5.4	27.2	1,985	4.0	27.8	1,389	1.41	20.68	0.700	0.19
Not FRPL	2,154	1.9	16.3	730	1.7	21.2	1,424	0.23	40.59	0.294	0.08
ELs	892	2.0	22.4	532	5.0	28.1	360	-2.97	37.76	0.057	-0.57
Non-ELs	6,411	4.4	24.9	3,367	2.9	24.1	3,044	1.48	22.60	0.836	0.26
SPED	1,017	13.7	35.5	610	8.3	34.5	407	5.47	24.85	0.092	0.35
Non-SPED	6,287	3.1	21.6	3,289	2.7	22.7	2,998	0.39	26.65	0.273	0.08
Cohorts 1, 2, and 3											
Full Sample	18,904	6.1	†	7,374	6.4	†	11,530	-0.32	16.81	0.942	-0.03
Subgroup											
Female	9,039	4.2	†	3,616	3.8	†	5,423	0.36	19.81	0.340	0.06
Male	9,652	7.5	†	3,755	9.1	†	5,897	-1.57	17.51	0.286	-0.13
Students of Color	11,703	6.7	†	4,249	6.5	†	7,454	0.21	19.44	0.925	0.02
White	6,811	4.8	†	3,113	6.0	†	3,698	-1.18	12.58	0.951	-0.14
FRPL	9,106	8.2	†	3,672	7.7	†	5,434	0.52	15.64	0.336	0.04
Not FRPL	5,715	2.7	†	2,141	4.5	†	3,574	-1.74 *	17.61	0.044	-0.32
ELs	2,662	6.2	†	1,014	7.4	†	1,648	-1.21	19.37	0.079	-0.12
Non-ELs	15,476	5.6	†	6,197	5.6	†	9,279	-0.03	16.76	0.537	0.00
SPED	2,781	16.6	†	1,040	11.7	†	1,741	4.85 *	17.80	0.019	0.25
Non-SPED	15,773	4.4	†	6,319	5.2	†	9,454	-0.82	17.69	0.474	-0.11

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations.

Source: AIR calculations from school-provided administrative data.

Exhibit E23. Impacts on Persistence to 10th Grade (i.e., Percentage of Students Persisting to 10th Grade), by Cohort and Subgroup

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohort 1											
Full Sample	6,305	94.5	32.2	4,246	90.7	29.1	2,059	3.83	33.69	0.090	0.35
Subgroup											
Female	2,953	95.3	27.0	1,973	93.1	29.3	980	2.24	32.08	0.196	0.25
Male	3,168	95.6	28.0	2,089	90.5	28.9	1,079	5.12 *	42.05	0.050	0.50
Students of Color	3,417	93.9	30.0	2,213	88.8	31.6	1,204	5.13	48.98	0.172	0.41
White	2,547	96.6	23.4	1,692	96.1	25.0	855	0.59	45.02	0.710	0.10
FRPL	2,742	94.6	25.0	2,071	88.4	33.8	671	6.24 *	33.62	0.013	0.51
Not FRPL	1,908	97.8	19.5	1,244	96.6	24.1	664	1.19	33.62	0.186	0.27
ELs	423	98.5	21.3	252	88.8	34.2	171	9.66	158.64	0.185	1.28
Non-ELs	5,222	95.3	26.7	3,339	93.3	28.2	1,883	2.06	34.71	0.263	0.24
SPED	899	95.8	31.7	550	91.0	33.2	349	4.81 *	35.72	0.023	0.49
Non-SPED	4,698	95.2	26.2	2,993	93.5	27.7	1,705	1.63	36.05	0.392	0.19
Cohort 2											
Full Sample	7,555	85.9	35.7	3,513	82.7	41.7	4,042	3.15	21.54	0.268	0.14
Subgroup											
Female	3,610	87.0	35.0	1,669	83.9	40.3	1,941	3.15	18.51	0.169	0.15
Male	3,929	85.2	36.1	1,837	82.1	42.8	2,092	3.11	26.57	0.391	0.14
Students of Color	6,016	83.9	36.9	2,948	79.8	43.2	3,068	4.18	21.07	0.180	0.17
White	1,505	94.6	25.9	554	91.8	34.6	951	2.79	47.75	0.349	0.27
FRPL	4,638	85.1	35.7	2,283	80.1	43.0	2,355	5.05	24.31	0.146	0.21
Not FRPL	1,889	92.5	31.5	994	90.0	36.3	895	2.49	19.14	0.100	0.19
ELs	1,959	79.7	40.6	821	77.2	44.6	1,138	2.46	31.80	0.647	0.09
Non-ELs	5,579	88.3	33.7	2,685	84.5	40.2	2,894	3.84 *	14.86	0.027	0.20
SPED	1,164	83.0	38.2	515	80.2	43.2	649	2.80	20.81	0.370	0.11
Non-SPED	6,373	86.7	35.1	2,990	83.5	41.3	3,383	3.18	21.94	0.251	0.15

(continues)

Exhibit E23. Impacts on Persistence to 10th Grade, by Cohort and Subgroup (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	13,860	88.4	†	7,759	85.0	†	6,101	3.35	18.15	0.065	0.18
Subgroup											
Female	6,563	89.1	†	3,642	86.2	†	2,921	2.93	16.03	0.066	0.17
Male	7,097	88.2	†	3,926	84.5	†	3,171	3.69	22.46	0.076	0.19
Students of Color	9,433	85.5	†	5,161	81.2	†	4,272	4.33	19.36	0.077	0.19
White	4,052	95.7	†	2,246	94.1	†	1,806	1.62	32.76	0.361	0.21
FRPL	7,380	88.4	†	4,354	82.9	†	3,026	5.46 **	19.70	0.008	0.28
Not FRPL	3,797	93.8	†	2,238	91.6	†	1,559	2.17 *	16.63	0.037	0.20
ELs	2,382	80.4	†	1,073	77.7	†	1,309	2.74	31.18	0.478	0.10
Non-ELs	10,801	89.4	†	6,024	85.8	†	4,777	3.57 *	13.66	0.013	0.20
SPED	2,063	86.2	†	1,065	82.9	†	998	3.31	17.98	0.055	0.16
Non-SPED	11,071	89.0	†	5,983	86.2	†	5,088	2.76	18.75	0.154	0.16

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations.

Source: AIR calculations from school-provided administrative data.

Exhibit E24. Impacts on Persistence to 10th Grade, by Cohort and Subgroup, *Compliance Adjusted*

		Treatment			Control						
	Sample Size	Mean	SD	*N*	Mean	SD	*N*	Difference	SE	*P*-value	Effect Size
					Cohort 1						
Full Sample	6,305	96.7	28.6	1,948	91.2	32.3	4,357	5.4	60.65	0.181	0.62
Subgroup											
Female	2,953	96.6	27.0	983	93.3	28.2	1,970	3.3	55.30	0.412	0.43
Male	3,168	97.9	30.0	964	91.1	27.5	2,204	6.8	78.14	0.120	0.93
Students of Color	3,417	99.0	32.4	875	85.7	29.9	2,542	13.3	203.58	0.397	1.69
White	2,547	96.8	24.4	1,067	96.1	23.5	1,480	0.7	55.14	0.381	0.12
FRPL	2,742	96.9	26.8	953	90.5	28.0	1,789	6.4	47.33	0.167	0.71
Not FRPL	1,908	98.1	20.0	768	96.8	22.0	1,140	1.3	40.58	0.233	0.33
ELs	423	100.0	28.4	80	79.6	27.4	343	20.4	784.08	0.377	6.30
Non-ELs	5,222	96.4	28.0	1,710	93.6	26.9	3,512	2.8	54.06	0.457	0.37
SPED	899	98.2	37.7	217	91.8	30.2	682	6.4	70.35	0.176	0.97
Non-SPED	4,698	96.0	26.8	1,638	93.7	26.7	3,060	2.3	55.71	0.587	0.29
					Cohort 2						
Full Sample	7,555	87.7	37.7	1,933	82.9	39.6	5,622	4.7	34.25	0.298	0.23
Subgroup											
Female	3,610	88.9	37.2	906	84.0	38.3	2,704	4.8	30.16	0.245	0.25
Male	3,929	86.9	38.2	1,027	82.3	40.6	2,902	4.6	41.15	0.373	0.21
Students of Color	6,016	86.2	40.0	1,480	80.3	40.6	4,536	6.0	32.24	0.142	0.26
White	1,505	96.8	26.5	449	90.3	33.7	1,056	6.4	124.92	0.306	0.71
FRPL	4,638	87.8	37.3	1,047	80.9	40.5	3,591	6.9	36.41	0.065	0.32
Not FRPL	1,889	94.0	34.2	657	89.7	33.8	1,232	4.3	35.76	0.481	0.36
ELs	1,959	80.5	43.3	584	77.3	43.0	1,375	3.2	41.43	0.532	0.12
Non-ELs	5,579	90.9	34.5	1,349	84.6	38.2	4,230	6.3 *	27.07	0.040	0.36
SPED	1,164	85.0	39.0	280	80.2	41.9	884	4.8	37.38	0.338	0.20
Non-SPED	6,373	88.4	37.5	1,652	83.7	39.1	4,721	4.7	34.19	0.313	0.24

(continues)

Exhibit E24. Impacts on Persistence to 10th Grade, by Cohort and Subgroup, *Compliance Adjusted* (continued)

	Sample Size	Treatment			Control			Difference	SE	*P*-value	Effect Size
		Mean	SD	*N*	Mean	SD	*N*				
Cohorts 1 and 2											
Full Sample	13,860	89.8	†	3,881	84.9	†	9,979	4.90	29.82	0.072	0.28
Subgroup											
Female	6,563	90.6	†	1,889	86.2	†	4,674	4.46	26.48	0.068	0.27
Male	7,097	89.3	†	1,991	84.2	†	5,106	5.07	36.41	0.094	0.28
Students of Color	9,433	86.6	†	2,355	80.4	†	7,078	6.15	31.85	0.124	0.27
White	4,052	96.8	†	1,516	95.2	†	2,536	1.63	50.45	0.473	0.27
FRPL	7,380	91.2	†	2,000	84.5	†	5,380	6.70 **	28.86	0.008	0.40
Not FRPL	3,797	95.8	†	1,425	92.8	†	2,372	2.99 *	26.83	0.035	0.35
ELs	2,382	80.5	†	664	77.3	†	1,718	3.20	41.37	0.598	0.12
Non-ELs	10,801	92.0	†	3,059	86.4	†	7,742	5.59 *	24.21	0.013	0.36
SPED	2,063	87.9	†	497	82.8	†	1,566	5.14	33.01	0.063	0.26
Non-SPED	11,071	90.5	†	3,290	86.5	†	7,781	4.02	29.14	0.154	0.24

Note: FRPL = free or reduced-price lunch; EL = English learner; SPED = special education. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. † The meta-analyses used to calculate impact estimates across cohorts do not allow us to report standard deviations.

Source: AIR calculations from school-provided administrative data.

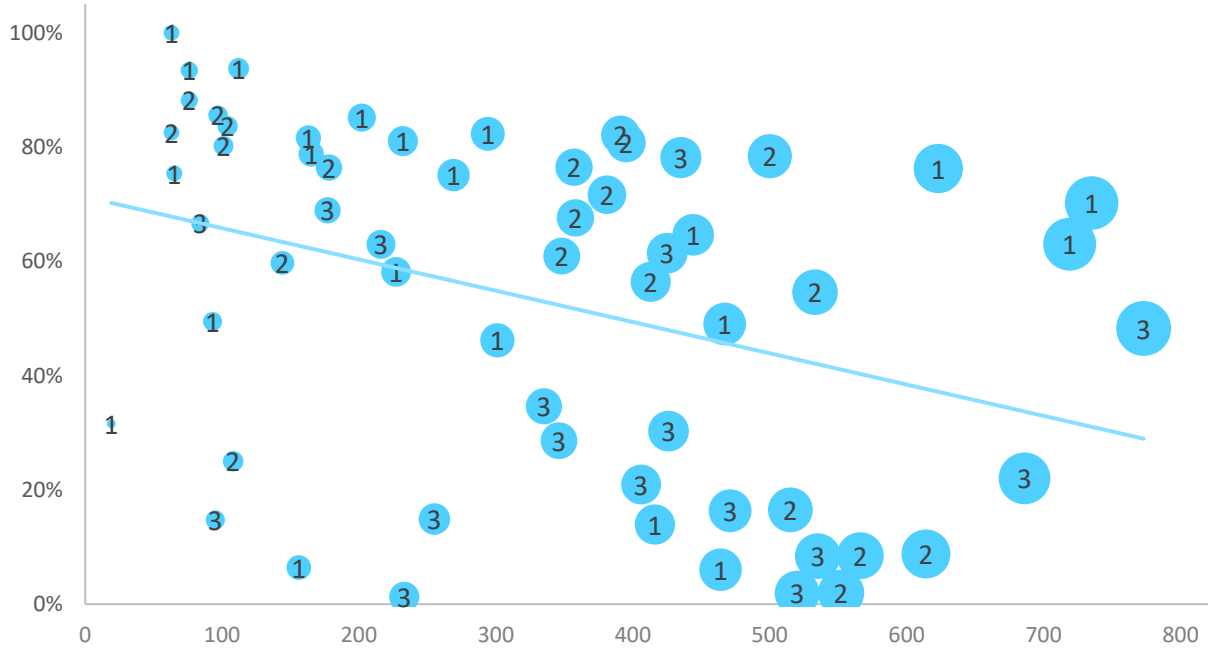
Appendix F. Student Survey Analytical Challenges

We administered a survey to the students in all study schools in the spring of ninth grade for cohorts 1 and 2 and in the fall of tenth grade for Cohort 3. This survey intended to measure the impact of BARR on student experiences and student engagement, like similar surveys in the two earlier impact evaluations of BARR (Bos et al., 2019; Corsello & Sharma, 2015). The survey measured student experiences and attitudes along the following six constructs: supportive relationships, expectations & rigor, engagement, emotional safety & school climate, future orientation, and self-efficacy. We used Rasch modeling to create the survey scales, validated the variable-creation process, and confirmed that the resulting scales were reliable (Cronbach's alpha ranging from 0.78 to 0.93).

This appendix describes analytical challenges we encountered during the survey administration and explains our decision not to include results from the analysis of survey data in the body of this report because we did not believe the data to be an accurate and reliable representation of the experiences and engagement of the students in this study.

The administration of the student survey was challenging for a variety of reasons. For cohorts 1 and 2, we were not always able to administer the survey in person during our planned site visits to study schools. As a result, there was variation in the timing of survey administration and in response rates across schools. In Exhibit F1, we show how survey response rates varied with school size. In addition to showing school size on the x-axis, we also varied the size of the school-level data points themselves to reflect their relative size.

Exhibit F1. Student Survey Response Rate by Student Sample Size in Schools That Administered the Survey (Numbers in Circles Indicate Cohort)



The trend line indicates that larger schools tended to have lower survey response rates, which exacerbates the consequences of survey nonresponse in cluster randomized controlled trials (RCTs) like these. The exhibit also shows that response rates were generally lowest for Cohort 3 schools, where most students completed the survey virtually due to the COVID-19 campus closures.

Exhibit F2 shows response rates by cohort and by treatment group. The survey response rates were similar in the two research groups, but this balance varied by cohort. In Cohort 1, control group students were more likely to respond to the survey than treatment group students, and in cohorts 2 and 3, treatment group students were more likely to respond.

Exhibit F2. Student Survey Response Rate by Cohort and Treatment Status in Schools That Administered the Survey

	Treatment	Control
Cohort 1	53%	71%
Cohort 2	55%	46%
Cohort 3	37%	22%
All cohorts	48%	46%

The relatively low overall response rates and the fact that they varied across research groups and cohorts raises concern about possible survey response bias. Such bias happens when survey respondents differ from nonrespondents in meaningful and statistically significant ways. Because these differences cannot always be controlled for, survey nonresponse can compromise the experimental study design by introducing selection bias into impact estimates (Shadish et al., 2002). The risk for such bias is magnified when survey response patterns differ between the research groups.

A common way to address potential survey nonresponse bias is to statistically control for measurable differences in student background characteristics between survey respondents and nonrespondents (or to reweight the sample using these characteristics). In the case of this evaluation, this was not possible because we were unable to link survey data to other student background or outcome data for the large majority (83 percent) of the students in the study sample. This is because the student identifiers were masked in many of the administrative data files we received for the study. Because of this, we could not use statistical controls or imputation strategies to reduce the effects of survey nonresponse on the impact estimates. We did not know enough about the respondents and nonrespondents to implement those strategies.

To get a better sense of the possible extent of survey nonresponse bias, we took the survey data from 21 schools that did not mask student ID variables in the administrative data they sent us and merged these administrative data with the student survey files. This resulted in a merged sample of approximately 6,000 students for whom we had administrative data and a survey response status. The survey response rate for these students was 58 percent in the treatment group and 62 percent in the control group. Exhibit F3 compares the baseline demographic characteristics of the survey respondents and nonrespondents in this sample. It shows that survey respondents were significantly more likely to be female and significantly less likely to be students of color, eligible for free or reduced-price lunch, English learners, or in special education. These differences were similar across the research groups, except for the gender variable. Survey respondents in the control group were more likely to be female than survey respondents in the treatment group. This was especially noticeable in Cohort 1 (not shown).

Exhibit F3. Student Demographic Background Characteristics by Student Survey Response Status and Treatment Status

Outcome	*N*	Resp.	Nonresp.	Difference	*P*-value	Standardized Difference
			Treatment Group			
Female (%)	3,818	50.8	47.9	2.95	0.075	0.07
Students of color (%)	3,669	50.2	57.4	-7.20***	0.000	-0.18
Free or reduced-price lunch (%)	3,774	57.0	63.2	-6.21***	0.000	-0.16
English learners (%)	3,777	5.7	9.7	-4.01***	0.000	-0.35
Special education (%)	3,774	10.6	22.7	-12.17***	0.000	-0.55
			Control Group			
Female (%)	2,231	51.6	42.5	9.14***	0.000	0.22
Students of color (%)	2,230	62.1	70.0	-7.89***	0.000	-0.21
Free or reduced-price lunch (%)	1,148	66.2	75.2	-9.01**	0.001	-0.27
English learners (%)	2,225	10.5	15.8	-5.22***	0.001	-0.28
Special education (%)	2,225	11.5	25.6	-14.16***	0.000	-0.59
			Full Sample			
Female (%)	6,049	51.1	45.9	5.21***	0.000	0.13
Students of color (%)	5,899	54.7	62.1	-7.33***	0.000	-0.18
Free or reduced-price lunch (%)	4,922	59.2	65.9	-6.68***	0.000	-0.17
English learners (%)	6,002	7.5	12.0	-4.45***	0.000	-0.31
Special education (%)	5,999	10.9	23.8	-12.90***	0.000	-0.57

Note: These data are from an unrepresentative subset of 21 schools for which we could match survey response status to administrative school data. They do not describe the study sample and should only be used for the purpose of the survey nonresponse analysis presented in Appendix F of the report.

Next, we compared respondents to nonrespondents using administrative outcome variables for which we presented impact estimates in Chapter 3. If anything, the differences between respondents and nonrespondents shown in Exhibit F4 were even greater than the differences we found in their demographic characteristics. Survey respondents had significantly better academic outcomes than their peers who did not respond to the survey. As expected, the difference was especially pronounced for chronic absenteeism: only 10.9 percent of survey respondents were chronically absent compared with 28.5 percent of nonrespondents. There were no major differences between the treatment and control groups in the extent of this apparent nonresponse bias, except for the PSAT scores, where the difference between respondents and nonrespondents was almost twice as large in the control group than in the treatment group. However, remember that treatment-control comparisons in Exhibits F3 and F4 are not from a representative sample of study schools and do not necessarily show valid comparisons of characteristics and outcomes between the treatment and control groups.

Exhibit F4. Student Academic Outcomes by Student Survey Response Status and Treatment Status

Outcome	*N*	Resp.	Nonresp.	Difference	*P*-value	Standardized Difference
Treatment Group						
Percent of core credits earned	3,846	92.1	80.2	11.88***	0.000	0.49
Passed all courses (%)	3,846	76.9	63.0	13.92***	0.000	0.41
GPA	3,743	2.8	2.3	0.50***	0.000	0.49
PSAT Total	2,752	874.1	848.3	25.72***	0.000	0.17
Chronically absent	3,992	10.6	27.7	-17.13***	0.000	-0.71
Ever suspended	4,000	9.5	16.4	-6.98***	0.000	-0.38
Persisted until 10th grade	4,000	93.2	76.7	16.42***	0.000	0.86
Control Group						
Percent of core credits earned	2,142	88.2	76.5	11.79***	0.000	0.45
Passed all courses (%)	2,142	69.4	50.4	19.04***	0.000	0.49
GPA	2,048	2.6	2.0	0.56***	0.000	0.53
PSAT Total	1,228	847.0	795.8	51.22***	0.000	0.36
Chronically absent	2,231	11.3	29.8	-18.56***	0.000	-0.73
Ever suspended	2,231	7.4	14.6	-7.21***	0.000	-0.46
Persisted until 10th grade	2,231	96.1	79.9	16.27***	0.000	1.11
Full Sample						
Percent of core credits earned	5,988	90.7	79.0	11.71***	0.000	0.47
Passed all courses (%)	5,988	74.1	58.8	15.37***	0.000	0.42
GPA	5,791	2.7	2.2	0.52***	0.000	0.50
PSAT Total	3,980	865.4	833.8	31.57***	0.000	0.21
Chronically absent	6,223	10.9	28.5	-17.60***	0.000	-0.72
Ever suspended	6,231	8.7	15.8	-7.13***	0.000	-0.41
Persisted until 10th grade	6,231	94.3	77.8	16.47***	0.000	0.94

Note: These data are from an unrepresentative subset of 21 schools for which we could match survey response status to administrative school data. They do not describe outcomes for the full study sample and should only be used for the purpose of the survey nonresponse analysis presented in Appendix F of the report.

One reason why we conducted the survey nonresponse analyses presented here is that preliminary student survey impact analyses produced anomalous results. These preliminary impact estimates, shown in Exhibit F5, were strongly and consistently negative for Cohort 1, mostly positive for Cohort 2, and mostly insignificant for Cohort 3. (Note that the estimates presented in Exhibit F5 include *all* student survey respondents, not just those from schools where we had access to student background data, such as those presented in Exhibits F3 and F4). Given that other data sources and outcome measures showed favorable estimated program effects, we became suspicious of the survey data.

Exhibit F5. Preliminary Impact Estimates on Student Survey Outcomes for All Student Survey Respondents (Without Controlling for Survey Nonresponse Bias)

	N	BARR	Control	Difference	*P*-value	Effect Size
Cohort 1						
Emotional safety and school climate	3,632	47.5	50.3	-2.85 **	0.001	-0.31
Expectations and rigor	3,674	47.4	49.4	-2.00 **	0.008	-0.21
Future orientation	3,633	48.0	50.5	-2.48 **	0.001	-0.26
Self-efficacy	3,608	50.2	52.5	-2.28 **	0.002	-0.23
Student engagement	3,656	47.9	50.5	-2.62 ***	0.000	-0.27
Supportive relationships	3,686	48.0	49.9	-1.86 **	0.001	-0.19
Cohort 2						
Emotional safety and school climate	3,175	49.9	49.4	0.48	0.478	0.05
Expectations and rigor	3,291	50.8	49.9	0.90 *	0.013	0.09
Future orientation	3,173	50.8	49.4	1.41 ***	0.000	0.13
Self-efficacy	3,125	49.5	48.7	0.77	0.150	0.08
Student engagement	3,241	50.7	50.1	0.56	0.176	0.05
Supportive relationships	3,368	51.4	49.8	1.65 ***	0.000	0.16
Cohort 3						
Emotional safety and school climate	1,868	52.8	52.5	0.39	0.786	0.04
Expectations and rigor	1,905	52.1	51.9	0.17	0.775	0.02
Future orientation	1,855	51.4	50.6	0.78	0.451	0.08
Self-efficacy	1,840	48.7	50.2	-1.46 **	0.006	-0.15
Student engagement	1,877	50.8	51.0	-0.11	0.843	-0.01
Supportive relationships	1,924	51.0	50.1	0.84	0.472	0.08
Cohorts 1, 2, and 3						
Emotional safety and school climate	8,675	49.4	50.1	-0.66	0.184	-0.07
Expectations and rigor	8,870	50.6	50.3	0.32	0.263	0.03
Future orientation	8,661	50.4	49.7	0.75 **	0.008	0.07
Self-efficacy	8,573	49.3	50.1	-0.75 *	0.025	-0.08
Student engagement	8,774	49.7	50.5	-0.73 **	0.006	-0.07
Supportive relationships	8,978	50.3	49.8	0.50	0.127	0.05

Note: These impact estimates are subject to nonresponse bias. They do not represent valid estimates of the impact of BARR on these outcomes. * = statistically significant at the $p < .05$ level; ** = statistically significant at the $p < .01$ level; *** = statistically significant at the $p < .001$ level. Combined Cohort 1, 2, and 3 results were obtained using a meta-analysis model that averaged the estimated difference across cohorts, giving more weight to estimates with greater precision. For additional details, see Appendix A.

Source: American Institutes for Research (AIR) calculations from AIR-administered student surveys. Scales are standardized to a mean of 50 and standard deviation of 10.

As a first step to understanding the preliminary impact results, we decided to test our prior research about the relationships between the student survey constructs and the other outcome variables. Prior research—including a mediation analysis in Bos et al., 2019—suggested that favorable effects on student experiences, relationships, and engagement mediate the effects of BARR on credit attainment, academic skills, and behavioral outcomes (including attendance and suspensions). That is, the correlation between BARR's impact on the student survey variables and on academic outcomes should be positive, which is contrary to what the preliminary student survey impact estimates in Exhibit F5 suggest.

Exhibit F6 shows the results of a correlation analysis, which again uses the sample for which we could match student survey data to administrative academic and behavioral outcomes. This correlation matrix provides support for the hypothesized theory of change underlying BARR and is consistent with our prior research. All correlations between survey variables and academic and behavioral outcome variables were in the expected direction, and most were statistically significant. These results were consistent across cohorts and research groups.

Exhibit F6. Correlation Between Academic Outcomes and Student Survey Outcomes for Students with Matched Data

	Supportive Relationships	Expectations & Rigor	Student Engagement	Emotional Safety	School Climate	Future Orientation	Self-Efficacy
Credits earned	**0.157**	**0.162**	**0.162**	**0.134**	**0.110**	**0.100**	**0.175**
Passed all core courses	**0.143**	**0.149**	**0.149**	**0.141**	**0.117**	**0.099**	**0.168**
GPA	**0.201**	**0.199**	**0.229**	**0.174**	**0.161**	**0.117**	**0.255**
PSAT Total	**0.095**	**0.087**	**0.087**	**0.089**	**0.096**	0.017	**0.129**
PSAT Math	**0.095**	**0.085**	**0.082**	**0.111**	**0.107**	**0.026**	**0.132**
PSAT ELA	**0.083**	**0.077**	**0.080**	**0.057**	**0.074**	0.008	**0.110**
Chronic absenteeism	-0.010	**-0.035**	-0.021	**-0.081**	**-0.067**	**-0.057**	**-0.060**
Ever suspended	**-0.078**	**-0.111**	**-0.073**	**-0.082**	**-0.121**	**-0.082**	-0.027
Persistence to 10th grade	-0.001	0.004	**0.050**	**0.070**	0.024	**0.055**	**0.072**

Note: Correlations in **boldface** were statistically significant at the 0.05 level. These data are from an unrepresentative subset of 21 schools for which we could match survey response status to administrative school data. They do not describe outcomes for the full study sample and should only be used for the purpose of the survey nonresponse analysis presented in Appendix F of the report.
Source: American Institutes for Research (AIR) calculations from AIR-administered student surveys and school- and district-provided administrative records.

The fact that these survey variables were positively correlated with academic outcomes is difficult to reconcile with an impact story in which program effects on student experiences are negative and program effects on academic outcomes are positive.[65]

The low student survey response rates, the degree to which they varied across schools, the evidence of nonresponse bias, and the fact that this bias varied by research group all undermine the validity of the preliminary student survey impact estimates. Together with the correlations presented in Exhibit F6, they have led us to conclude that these preliminary student survey impact estimates were not a valid representation of BARR's impact on student experiences and that the student survey data are too flawed to use them for the purpose of impact analyses for this report. We therefore did not include or reference these impact estimates in the body of this report or in the executive summary.

[65] We re-estimated the impact estimates using only the sample for whom we had matched data to make sure that the anomalous impact estimates also appeared in this matched subsample. They did.

Appendix G. Scale-Up Measurement

Exhibit G1 provides details on the BARR scale-up strategies measured during the study as well as findings. These findings align with the progress reported in Chapter 4 and summarized in Exhibit 4.1.

Exhibit G1. BARR Scale-Up Measurement and Findings

Indicators	Data Sources	Threshold	Year 1 (2017–18 School Year) Finding	Year 2 (2018–19 School Year) Finding	Year 3 (2019–20 School Year) Finding	Year 4 (2020–21 School Year) Finding
Strategies for Expanding the BARR Model						
Key Strategy: Fortify BARR Center Infrastructure						
a. BARR Center maintains staff-to-school ratio.	BARR Center staffing data	Year 1: n/a Years 2–4: BARR staff-to-school ratios do not exceed staffing expectations. The expectations were that coaches oversaw 12 schools, regional managers oversaw 10 schools, and program managers oversaw five schools.	n/a	BARR staffing expectations were met.	BARR staffing expectations were not met. One regional manager was responsible for 11 schools.	BARR staffing expectations were met.
b. BARR Center fills vacancies and open positions.	BARR Center staffing data	Year 1: n/a Years 2–4: 67% of vacancies are filled.	n/a	100% of vacancies were filled.	100% of vacancies were filled.	100% of vacancies were filled.
c. BARR Center is fiscally sustainable.	IRS certification letter and financial statements (990 series)	Year 1: n/a Year 2: BARR receives 501(c)(3) status. Years 3–4: BARR is fiscally sustainable (990 series).	n/a	BARR received 501(c)(3) status (n=1, BARR national).	BARR was fiscally sustainable based on financial statements (990 series) (n=1, BARR national).	BARR was fiscally sustainable based on financial statements (990 series) (n=1, BARR national).

Indicators	Data Sources	Threshold	Year 1 (2017–18 School Year) Finding	Year 2 (2018–19 School Year) Finding	Year 3 (2019–20 School Year) Finding	Year 4 (2020–21 School Year) Finding
d. BARR Center builds BARR model visibility.	Communications plan and copies of relevant (1) presentations, (2) local media articles or interviews, (3) social media records, (4) website updates, and/or (5) partnerships with national organizations	Year 1: n/a Years 2–4: BARR will have met four of the five (80%) annual benchmarks set for increasing program visibility (presentations, articles/interviews, social media records, website updates, and/or signed partnerships with national organizations).	n/a	BARR met four of the five (80%) annual benchmarks set for increasing program visibility.	BARR met five of the five (100%) annual benchmarks set for increasing program visibility.	BARR met four of the five (80%) annual benchmarks set for increasing program visibility.
Key Strategy: Expand School and District Awareness						
a. BARR Center shares school communication strategies.	Communication Strategies document delivery records	Year 1: n/a Years 2–4: 84% of schools that have implemented for 1 year or more receive the school Communication Strategies document.	Did not have adequate data to conduct analysis.	Did not have adequate data to conduct analysis.	Did not have adequate data to conduct analysis.	Did not have adequate data to conduct analysis.
b. Schools use BARR communication strategies.	School web metrics OR self-report (e.g., tweets, posts, news articles)	Year 1: n/a Years 2–4: 51% of schools that have implemented for 1 year use BARR communication strategies.	n/a	13.3% of schools that have implemented for 1 year used BARR communication strategies (2 of 15 schools [CT1 T]).	18.2% of schools that have implemented for 1 year used BARR communication strategies (6 of 33 schools [CT1 T&C + CT2 T]).	20.0% of schools that have implemented for 1 year used BARR communication strategies (11 of 55 schools [W1 T&C + CT2 T&C + CT3 T]).

Indicators	Data Sources	Threshold	Year 1 (2017–18 School Year) Finding	Year 2 (2018–19 School Year) Finding	Year 3 (2019–20 School Year) Finding	Year 4 (2020–21 School Year) Finding
Strategies for Sustaining the BARR Model						
Key Strategy: Sustain Cost of Services						
a. BARR Center shares cost of services materials.	Cost of Services document delivery records	Years 1–2: n/a Years 3–4: 100% of schools that are in their third year of implementation (and did not discontinue participation in previous years) receive the Cost of Services document.	n/a	n/a	100% of schools that are in their third year of implementation received the Cost of Services document (11 of 11 CT1 T schools that were still implementing in their third year).	96.4% of schools that are in their third or more year of implementation (and did not discontinue participation in Year 3) received the Cost of Services document (27 of 28 CT1 T&C + CT2 T schools that were still implementing in their third year).
b. Schools in need of supplemental funding secure funds for the subsequent year.	School OR regional hub self-report on need for additional funding and funding pursuits	Years 1–2: n/a Years 3–4: 34% of schools in need of supplemental funding in their third or more year of BARR implementation secure funding to help sustain BARR beyond the randomized controlled trial in the subsequent year.	n/a	n/a	71.4% of schools in need of supplemental funding in their third year of BARR implementation secured philanthropic or third-party funding to help sustain BARR beyond the RCT in the subsequent year (5 of 7 schools [CT1 T]).	42.1% of schools in need of supplemental funding in their third or more year of BARR implementation secured philanthropic or third-party funding to help sustain BARR beyond the RCT in the subsequent year (8 of 19 schools [CT1 T&C + CT2 T]).

Indicators	Data Sources	Threshold	Year 1 (2017–18 School Year) Finding	Year 2 (2018–19 School Year) Finding	Year 3 (2019–20 School Year) Finding	Year 4 (2020–21 School Year) Finding
Key Strategy: Sustain Coach Training At Scale						
a. New and continuing coaches complete BARR training.	Attendance records for coach trainings	Years 1–4: 84% of coaches complete minimum BARR training requirements.	89.5% of coaches completed minimum BARR training requirements (17 of 19 regional managers and coaches).	90.3% of coaches completed minimum BARR training requirements (28 of 31 regional managers and coaches).	94.4% of coaches completed minimum BARR training requirements (34 of 36 regional managers and coaches).	95.0% of coaches completed minimum BARR training requirements (38 of 40 regional managers and coaches).
b. BARR Center creates coach mastery rubric.	BARR Center coach mastery rubric development records and conversations with BARR staff	Year 1: n/a Year 2: BARR documents coaching expectations (Y/N). Year 3: BARR examines proof of concept of coach rubric (Y/N). Year 4: BARR pilots rubric with BARR coaches (Y/N).	n/a	BARR drafted the coach rubric.	BARR held working meetings and refined the coach rubric.	BARR introduced the coaching rubric to BARR coaches in spring 2021.
Key Strategy: Sustain School-Level Staff Training At Scale						
a. Educators complete BARR training.	BARR records of number of school staff at each school training	Years 1–4: 85% of schools have at least three school staff members complete BARR training.	93.3% of schools had at least three school staff members complete BARR training (14 of 15 schools [CT1 T]).	97.0% of schools had at least three school staff members complete BARR training (32 of 33 schools [CT1 T&C + CT2 T]).	96.4% of schools had at least three school staff members complete BARR training (53 of 55 schools [CT1 T&C + CT2 T&C + CT3 T]).	95.5% of schools had at least three school staff members complete BARR training (63 of 66 schools [CT1 T&C + CT2 T&C + CT3 T&C]).
b. Schools use BARR database.	BARR database reporting completion records	Year 1: n/a Years 2–4: 85% of schools import data.	n/a	40.9% of schools imported data (9 of 22 schools [CT1 T&C]).	56.8% of schools imported data (25 of 44 schools [CT1 T&C + CT2 T&C]).	39.4% of schools imported data (26 of 66 schools [CT1 T&C + CT2 T&C + CT3 T&C]).

Indicators	Data Sources	Threshold	Year 1 (2017–18 School Year) Finding	Year 2 (2018–19 School Year) Finding	Year 3 (2019–20 School Year) Finding	Year 4 (2020–21 School Year) Finding
Key Strategy: Sustain School-Level Implementation						
a. Schools implement BARR model with fidelity.	School-level fidelity analysis using scores collected from BARR coaches	Years 1–4: 67% of schools receive an adequate rating on four out of five priority components and two out of three additional components.	33% of schools received an adequate rating on four out of five priority components and two out of three additional components (5 of 15 schools [CT1 T]).	55% of schools received an adequate rating on four out of five priority components and two out of three additional components (18 of 33 schools [CT1 T&C + CT2 T]).	29% of schools received an adequate rating on four out of five priority components and two out of three additional components (16 of 55 schools [CT1 T&C + CT2 T&C + CT3 T]).	36% of schools[66] received an adequate rating on four out of five priority components and two out of three additional components (21 of 58 schools[67] [CT1 T&C + CT2 T&C + CT3 T&C]).
b. BARR Center creates school accreditation process.	BARR Center accreditation development records and conversations	Years 1–2: n/a Year 3: BARR drafts school accreditation process (Y/N). Year 4: BARR examines proof of concept of school accreditation (Y/N).	n/a	n/a	BARR drafted the school accreditation process.	BARR refined the school accreditation process.

Note: CT = Cohort, T = Treatment, C = Control

[66] BARR coaches were unable to measure implementation of one or more BARR components at 12 schools during Year 4 due to COVID-19. In these cases, implementation was rated as a "1," which may likely be lower than actual implementation.

[67] From our original sample of 66 study schools, eight Cohort 1 treatment schools did not secure funding for a fourth year of BARR implementation.