



# Building Assets and Reducing Risks (BARR) Validation Study Final Report

APRIL 2019

Johannes M. Bos | Sonica Dhillon | Trisha Borman

MAKING RESEARCH RELEVANT



# Building Assets and Reducing Risks (BARR) Validation Study

## Final Report

APRIL 2019

Johannes M. Bos | Sonica Dhillon | Trisha Borman

*With Brenna O'Brien, Cheryl Graczewski, So Jung Park, Feng Liu,  
Ethan Adelman-Sil, Lynn Hu*



AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW  
Washington, DC 20007-3835  
202.403.5000

[www.air.org](http://www.air.org)

Copyright © 2019 American Institutes for Research. All rights reserved.

## **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, providing us with responses to our surveys, participating in reading and mathematics 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 eleven 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 communications with parents and other stakeholders in their districts and communities. Strong partnerships with schools are critical to the success of education evaluations, and these eleven 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, has been a major driver of innovation and rigorous effectiveness research in the U.S. education system. The i3 validation grant awarded to the BARR Center enabled BARR to implement its program in the eleven study schools in this evaluation and in 35 additional “dissemination” schools throughout the United States. The grant also paid for all of our data collection, analysis, and reporting and enabled us to present our findings at national BARR conferences and other venues.

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, our quality assurance reviewer, and Harrison Greene, our technical project reviewer. We also want to acknowledge our team members Brenna O’Brien, Megan Brown, Connie Chandra, David Gorsky, Amelia Auchstetter, Grace Fleming, and Thomas Fink.

Last, we want to thank the BARR Center and its staff and leadership, including Angela Jerabek, the director and originator of BARR; Justin Barbeau; Maryann Corsello; Susan Savell; and Rob Metz. They were genuine partners throughout this evaluation of their program, and they provided us with invaluable program data and insights. Most importantly, they successfully recruited and retained eleven schools for the evaluation. Their support was critical to maintaining the schools’ engagement, compliance with study protocols, and collaboration with data collection.

# Contents

	<b>Page</b>
Acknowledgments.....	iv
Executive Summary.....	1
Overview of the BARR Model and Evaluation Design.....	7
What Problem Does BARR Address? .....	7
Focus of the BARR Model .....	8
Development of the BARR Model.....	9
Components of the BARR Model .....	10
Evaluation Design.....	13
Study Design .....	13
Measurement and Data Collection.....	15
Baseline Characteristics of the Study Schools and Their Students.....	19
Did Study Schools Implement BARR as Intended? .....	21
To What Extent Were the Eight BARR Strategies Implemented as Intended? .....	22
Implications of the Implementation Findings and Study Design for the Treatment Contrast.....	22
What Were the Facilitators of and Barriers to Successful Implementation?.....	25
How Could Implementation Be Improved? .....	27
Did BARR Change Teacher Experiences and Teacher Practice? .....	28
Did BARR Impact Student Achievement, Experience, and Behaviors?.....	31
Impacts on Student Achievement .....	31
Impacts on Student Experiences .....	38
Impacts on Student Behavior .....	40
How Did These Impacts Vary by School?.....	42
Study Limitations .....	48
Discussion and Conclusion .....	49
References .....	53
Appendix A. BARR Logic Model.....	58
Appendix B. Data Collection Instruments.....	59
Appendix C. Attrition Rates for Outcome Measures .....	94
Appendix D. Analysis of Baseline Equivalence.....	97

Appendix E. Fidelity Measurement.....	102
Appendix F. Statistical Models.....	107
Appendix G. NWEA Sensitivity Analyses.....	109
Appendix H. Supplemental Student Behavior Subgroup Analyses.....	112
Appendix I. Supplemental Impact Estimates by School .....	114
Appendix J. Supplemental Exploratory Impact Estimates by NWEA Baseline Test Score Quartile .....	120

## Exhibits

	Page
Exhibit ES1. Differences in Teacher Outcomes Between BARR and Control Teachers .....	3
Exhibit ES2. Summary of Student-Level Impacts .....	4
Exhibit ES3. Percentage of Students Failing at Least One Core Course in Ninth Grade .....	5
Exhibit 1. Analytic Sample by Data Source and Group .....	17
Exhibit 2. Characteristics of Schools in Cohorts 1, 2, and 3.....	19
Exhibit 3. Randomized Student Sample for Cohorts 1, 2, and 3.....	20
Exhibit 4. Reporting Fidelity for Each BARR Key Component: All Three Cohorts.....	23
Exhibit 5. Differences in Teacher Experiences in Participating Schools .....	29
Exhibit 6. Impacts on Students Failing One or More Core Courses (Full Sample and Subgroups) .....	32
Exhibit 7. Impacts on Students’ Grade Point Average in Core Courses (Full Sample and Subgroups) .....	33
Exhibit 8. Impacts on Standardized Reading and Mathematics Scale Scores and Percentage Meeting Growth Projections .....	34
Exhibit 9. Impacts on Standardized Reading Scale Scores, by Subgroup .....	35
Exhibit 10. Impacts on Reading Growth Projections Met, by Subgroup .....	35
Exhibit 11. Impacts on Standardized Mathematics Scale Scores, by Subgroup.....	36
Exhibit 12. Impacts on Mathematics Growth Projections Met, by Subgroup.....	37
Exhibit 13. Impacts on Standardized Reading and Mathematics Scale Scores and Growth Projections, by Cohort .....	38
Exhibit 14. Impacts on Student Experiences Measured by Student Surveys .....	39

Exhibit 15. Impacts on Student Experiences From Student Surveys by Subgroup (Effect Sizes Only) ..... 40

Exhibit 16. Impacts on Students’ Chronic Absenteeism, Suspensions, and Persistence to Grade 10 ..... 41

Exhibit 17. Estimated Effects on Students’ Failing One or More Core Courses, by School (11 Schools)..... 43

Exhibit 18. Estimated Effects on Students’ Reading NWEA Scores, by School (10 Schools) ..... 43

Exhibit 19. Estimated Effects on Students’ NWEA Mathematics Scores, by School (10 Schools)..... 44

Exhibit 20. Estimated Effects on Students’ Reading NWEA Growth Projections Met, by School (10 Schools) ..... 44

Exhibit 21. Estimated Effects on Students’ NWEA Mathematics Growth Projections Met, by School (10 Schools) ..... 45

Exhibit B1.1. Student Survey Constructs, Outcome Measures, Data Sources for Assessing the Impact of BARR on Student Experiences..... 59

Exhibit B2.1. Teacher Survey Constructs, Outcome Measures, and Data Sources for Assessing the Impact of BARR on Student Experiences ..... 67

Exhibit C1. Attrition for NWEA Reading Scores (10 schools) ..... 94

Exhibit C2. Attrition for NWEA Mathematics Scores (10 schools)..... 94

Exhibit C3. Attrition for Core Course Failure (11 schools) ..... 94

Exhibit C4. Attrition for Supportive Relationships (11 schools) ..... 95

Exhibit C5. Attrition for Expectations and Rigor (11 schools)..... 95

Exhibit C6. Attrition for Student Engagement (11 schools)..... 95

Exhibit C7. Attrition for Sense of Belonging (11 schools) ..... 95

Exhibit C8. Attrition for Social and Emotional Learning (11 schools) ..... 96

Exhibit C9. Attrition for Grit (11 schools)..... 96

Exhibit D1. Treatment and Comparison Group Characteristics at Baseline for NWEA Reading (10 Schools)..... 97

Exhibit D2. Treatment and Comparison Group Characteristics at Baseline for NWEA Mathematics Baseline Covariates (10 Schools) ..... 97

Exhibit D3. Treatment and Comparison Group Characteristics at Baseline for Course Failure (11 Schools)..... 98

Exhibit D4. Treatment and Comparison Group Characteristics at Baseline for Supportive Relationships (11 Schools) ..... 98

Exhibit D5. Treatment and Comparison Group Characteristics at Baseline for Expectation and Rigor (11 Schools) ..... 99

Exhibit D6. Treatment and Comparison Group Characteristics at Baseline for Student Engagement (11 Schools)..... 99

Exhibit D7. Treatment and Comparison Group Characteristics at Baseline for Sense of Belonging (11 Schools)..... 100

Exhibit D8. Treatment and Comparison Group Characteristics at Baseline for Social and Emotional Learning (11 Schools) ..... 100

Exhibit D9. Treatment and Comparison Group Characteristics at Baseline for Grit (11 Schools)..... 101

Exhibit E1. Observations Collected From Participating Schools..... 102

Exhibit E2. Fidelity Matrix ..... 104

Exhibit E3. Changes in Number of Indicators and Thresholds From Participating Schools..... 106

Exhibit F3.1. The Total Number (%) of Imputed Missing Cases for NWEA Reading (10 Schools)..... 108

Exhibit F3.2. The Total Number (%) of Imputed Missing Cases for NWEA Mathematics (10 Schools)..... 108

Exhibit F3.3. The Total Number (%) of Imputed Missing Cases for Course Failure (11 Schools)..... 108

Exhibit G1. Impacts on Standardized Reading Scale Scores by Subgroup (Five schools) ..... 109

Exhibit G2. Impacts on Standardized Mathematics Scale Scores by Subgroup (Five schools) ..... 109

Exhibit G3. Impacts on Reading Growth Projections Met by Subgroup (Five schools) ..... 110

Exhibit G4. Impacts on Mathematics Growth Projections Met by Subgroup (Five schools)..... 111

Exhibit H1. Impacts on Students’ Chronic Absenteeism by Subgroup (Nine schools) ..... 112

Exhibit H2. Impacts on Students’ Suspensions by Subgroup (10 schools) ..... 112

Exhibit H3. Impacts on Students’ Persistence to Grade 10 by Subgroup (11 schools)..... 113

Exhibit I1. Estimated Effects on Students’ Grade Point Average, by School (11 Schools) ..... 114

Exhibit I2. Estimated Effects on Students’ Expectations and Rigor, by School (11 schools) ..... 115

Exhibit I3. Estimated Effects on Students’ Engagement, by School (11 schools)..... 115



Exhibit I4. Estimated Effects on Students’ Supportive Relationships, by School (11 schools) ..... 116

Exhibit I5. Estimated Effects on Students’ Social and Emotional Learning, by School (11 Schools)..... 116

Exhibit I6. Estimated Effects on Students’ Sense of Belonging, by School (11 Schools) ..... 117

Exhibit I7. Estimated Effects on Students’ Grit, by School (11 Schools)..... 117

Exhibit I8. Estimated Effects on Students’ Suspensions, by School (Seven Schools) ..... 118

Exhibit I9. Estimated Effects on Students’ Chronic Absenteeism, by School (Nine Schools) ..... 118

Exhibit I10. Estimated Effects on Students’ Persistence to Grade 10, by School (11 Schools)..... 119

Exhibit J1. Impacts on Students Failing One or More Core Courses, by Quartile (11 Schools)..... 120

Exhibit J2. Impacts on Students’ Grade Point Average in Core Courses, by Quartile (11 Schools)..... 120

Exhibit J3. Impacts on Standardized Reading Scale Scores, by Quartile (10 schools) ..... 120

Exhibit J4. Impacts on Standardized Mathematics Scale Scores, by Quartile (10 schools)..... 121

Exhibit J5. Impacts on Reading Growth Projections Met, by Quartile (10 schools) ..... 121

Exhibit J6. Impacts on Mathematics Growth Projections Met, by Quartile (10 Schools) ..... 121

Exhibit J7. Impacts on Students’ Expectations and Rigor, by Quartile (11 Schools) ..... 122

Exhibit J8. Impacts on Students’ Engagement, by Quartile (11 Schools) ..... 122

Exhibit J9. Impacts on Students’ Supportive Relationships, by Quartile (11 Schools) ..... 122

Exhibit J10. Impacts on Students’ Social and Emotional Learning, by Quartile (11 Schools)..... 123

Exhibit J11. Impacts on Students’ Sense of Belonging, by Quartile (11 Schools)..... 123

Exhibit J12. Impacts on Students’ Grit, by Quartile (11 Schools) ..... 123

Exhibit J13. Impacts on Students’ Suspensions, by Subgroup (10 Schools) ..... 124

Exhibit J14. Impacts on Students’ Chronic Absenteeism, by Subgroup (Nine Schools) ..... 124

Exhibit J15. Impacts on Students’ Persistence to Grade 10, by Subgroup (11 Schools) ..... 124



## **Executive Summary**

Building Assets, Reducing Risks (BARR) is a comprehensive, strength-based approach that uses eight interlocking strategies to build intentional staff-to-staff, staff-to-student, and student-to-student relationships in secondary schools. On the basis of prior evidence of model effectiveness, the program developers expect that schools that implement the program with fidelity will see improvements in school climate, teacher experiences, student engagement, and, over time, academic outcomes. BARR is currently being implemented in more than 100 schools throughout the United States, ranging from large, urban high schools to small, rural middle schools and high schools.

The BARR model requires a 3-year commitment from participating schools. During this time, the BARR team works with these schools to provide professional development, coaching, the I-Time curriculum (a social and emotional curriculum), and administrative supports. Schools commit to organizing their core schedule into blocks/teams so that core subject teachers and their students get to know each other well and core subject teachers can more effectively collaborate and share information between staff. After organizing their master schedule to accommodate BARR, school staff commit to conducting regular block/team meetings and risk-review meetings to discuss the progress, challenges, and assets of all students in their blocks and to develop and implement interventions as necessary. The BARR team works with schools to improve the quality of these meetings and to help them implement other BARR activities, including the I-Time curriculum. Schools that have adopted the BARR model are part of an ongoing learning community beyond their 3-year commitment and participate in annual BARR conferences at which they share their BARR experiences and lessons learned.

## **About This Study**

This is the final report of a large-scale independent evaluation of the BARR model in ninth grade in eleven high schools in Maine, California, Minnesota, Kentucky, and Texas. This sample of schools included large and small schools in urban, suburban, and rural areas, serving students from a wide range of demographic and socio-economic backgrounds. Funded with a validation grant from the U.S. Department of Education's Investing in Innovation (i3) program and carried out by researchers at the American Institutes for Research (AIR), this evaluation used random assignment of ninth-grade students to BARR and control conditions to estimate the impacts of the BARR model after one year. The evaluation also assessed the fidelity of implementation of BARR in the eleven study schools and identified barriers to and facilitators of successful implementation. The evaluation focused on several teacher- and student-level outcomes. The teacher outcomes included measures of teacher collaboration, and use of data, among others. The academic outcomes included course failure, students' grade point average (GPA), and performance on the Northwest Evaluation Association's (NWEA) Measures of Academic

Progress (MAP) standardized reading and mathematics assessments. Student-reported experiences included measures of supportive relationships, perceptions of teachers' expectations of them, student engagement, and others. In addition to these outcomes, the report includes impact estimates for attendance, suspensions, and persistence into 10th grade.

More than 4,000 ninth-grade students participated in the evaluation. AIR randomly assigned students within each school to receive BARR supports during their ninth-grade year or to be in a business-as-usual control group. Participating schools agreed to limit specific BARR activities to the treatment group of students within their school to reduce the risk of contamination. The eleven participating study schools were distributed across three cohorts, each of which participated in the randomized controlled trial for a single school year. Cohort 1 included three schools participating in 2014–15, Cohort 2 included three schools participating in 2015–16, and Cohort 3 included five schools participating in 2016–17. This study presents results across all three of these cohorts after each cohort's first year of implementation.

In addition to the 11-school study that forms the backbone of this report, the i3 validation grant supported dissemination of the BARR model in 35 additional schools in the following five states: Kentucky, Maine, Minnesota, North Carolina, and Wisconsin. Where relevant, evidence from BARR's internal evaluation of this expansion effort is referenced throughout this report.

## **What Were We Looking For?**

From its first school in Minnesota in 1998, the driving purpose of the BARR model has been to reduce course failure and increase academic success for all students. To get there, BARR schools work to improve staff-to-staff, student-to-staff, and student-to-student relationships; to empower teachers; and to leverage the strengths and assets that each student brings to school. This evaluation examined differences in teacher experiences between BARR teachers and control teachers. Improvements in these teacher experiences may translate into subsequent differences in teacher effectiveness, which in turn may enhance the academic learning opportunities for students. Improved relationships between teachers and students and between students and their peers were also expected to translate into impacts on student experiences and attitudes, and into impacts on behavioral outcomes such as attendance and suspensions. All of these are possible precursors of subsequent impacts on course completion and other academic outcomes.

BARR aims to benefit disadvantaged subgroups of students and this large-scale impact evaluation was designed to include separate estimates of impacts for different subgroups and schools. These estimates make it possible to determine whether and how BARR influenced achievement/opportunity gaps between different subgroups (e.g., by gender or student socio-

economic status). These analyses also enabled us to document the consistency and pervasiveness of program benefits across a range of different school settings.

### What Did We Find?

Using a combination of qualitative and quantitative data, the evaluation found that the key components of the BARR model were **implemented with fidelity** in the eleven study schools. These schools had to overcome logistical challenges and time constraints to implement program activities with fidelity. We found that having dedicated BARR coordinators, strong leadership buy-in, and ongoing coaching from the BARR Center helped support consistent program implementation. Implementation fidelity also benefited from the apparent popularity of the BARR program and its components with teachers, school staff, and students, as reported in surveys and interviews. Teachers and students valued the perceived effects of BARR on staff-to-staff, student-to-staff, and student-to-student relationships in the schools.

The experiences and practices of BARR teachers differed from those of control teachers in several meaningful ways, as shown in Exhibit ES1. Because teachers were not randomly assigned to BARR or the control group, these differences cannot be conclusively attributed to BARR, but they are consistent with other ongoing research on the BARR model.

Compared with their counterparts in the control group, BARR teachers reported

- more positive views about their colleagues and about collaborating with them,
- that they were more satisfied with the supports available to them in their school,
- that they were more likely to use data to inform their instruction,
- higher levels of self-efficacy, and
- more positive views about their students.

#### Exhibit ES1. Differences in Teacher Outcomes Between BARR and Control Teachers

Outcome	BARR	Control	Difference
Teacher collaboration with and view of colleagues	56.68	45.02	11.67‡
Teacher use of data	55.70	45.68	10.03‡
View of the school’s supports	53.47	47.55	5.92†
Teacher self-efficacy	52.92	47.89	5.03†
Perception of students’ behavior	52.72	48.00	4.72*
View of students’ observed behavior	52.59	48.09	4.49*
Interaction with parents	52.07	48.54	3.54
View of student accountability	50.26	49.80	0.46

Source: AIR calculations from AIR-administered teacher surveys of 113 ninth-grade teachers (48 BARR teachers and 65 control teachers).

Note. \* = statistically significant at the  $p < .05$  level; † = statistically significant at the  $p < .01$  level; ‡ = statistically significant at the  $p < .001$  level.

The BARR model positively affected several different student outcomes, as summarized in Exhibit ES2. The differences between BARR and the control group shown in this exhibit are unbiased estimates of the effects of BARR. Below are the results.

- BARR significantly reduced course failure and increased students’ GPA at the end of ninth grade.
- BARR did not have statistically significant impacts on reading and mathematics achievement as measured with the NWEA MAP assessment administered at the end of ninth grade.
- A larger share of students in the BARR group met their NWEA MAP growth target at follow-up, but BARR did not have a statistically significant impact on this outcome.
- Relative to the control group, BARR significantly improved some aspects of student experiences in school, including supportive relationships, expectations and rigor, and student engagement.
- BARR did not have a statistically significant impact on attendance, suspensions, or persistence into 10th grade.

**Exhibit ES2. Summary of Student-Level Impacts**

Outcome	N	BARR	Control	Difference	Effect size
<b>Academic Outcomes</b>					
Failed at least one core course (%)	3,383	29.2	40.6	-11.5‡	0.31
Grade point average	3,376	2.58	2.48	0.10‡	0.11
NWEA reading scores	2,595	220	220	0.13	0.01
NWEA mathematics scores	2,715	228	228	0.16	0.01
NWEA reading growth met (%)	2,218	59.1	55.2	3.9	0.10
NWEA mathematics Growth met (%)	2,252	59.0	56.7	2.3	0.06
<b>Student-Reported Experiences</b>					
Supportive relationships	2,716	51.5	48.7	2.90‡	0.29
Expectations and rigor	2,745	51.3	48.9	2.48‡	0.25
Student engagement	2,721	50.6	49.5	1.06†	0.11
Sense of belonging	2,692	50.2	49.8	0.45	0.04
Social and emotional learning	2,700	50.1	49.9	0.16	0.02
Grit	2,685	50.0	50.0	-0.02	0.00
<b>Behavioral Outcomes</b>					
Chronic absence (%)	3,275	23.1	21.9	1.2	0.04
Suspensions (%)	3,806	6.5	6.7	-0.2	0.02

Outcome	N	BARR	Control	Difference	Effect size
Persistence to grade 10 (%) <sup>a</sup>	2,863	88.8	87.4	1.4	0.08

Source: AIR calculations from school-provided administrative data, school-administered NWEA assessments, and AIR-administered student surveys. Table ES 2 presents the results from a statistical comparison between BARR and control students on a range of Likert-style survey measures, in the form of scale scores.

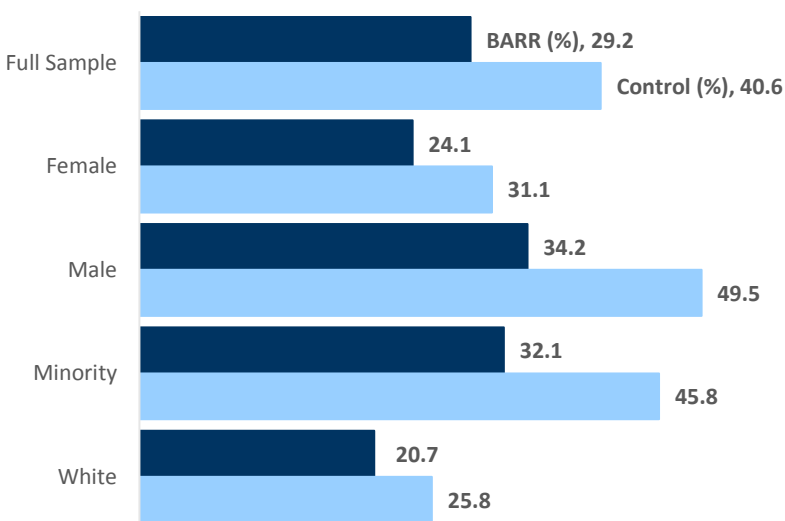
Note. † = statistically significant at the  $p < .01$  level; ‡ = statistically significant at the  $p < .001$  level.

<sup>a</sup> Due to administrative data constraints, this outcome is measured as enrolling in 10th grade at the same school. Because many students on the initial ninth-grade rosters never formally enrolled, the sample for this outcome only includes students who were enrolled in the study school in ninth grade for at least 20 days.

The BARR model positively affected the course failure outcome for most subgroups of students, as shown in Exhibit ES3. Similar subgroup analyses for other outcomes are included in the body of the report. Because the patterns of subgroup impacts were mostly similar across outcomes, only the course failure subgroup analysis is presented in detail in this Executive Summary.

BARR’s positive impact on course failure was strongest for males and minority students, demographic subgroups more at risk of course failure overall. The larger positive impacts for males and minority students, compared to females and white students respectively, indicate that BARR reduced performance gaps for course failure. For example, the difference in the course failure rate between white students and students of color was 20.0 percentage points in the control group but only 11.4 percentage points in the BARR group. This represents a reduction of the performance gap for students of color by almost half for this outcome. Other subgroups for whom BARR had generally stronger impacts on other outcomes include students eligible for free or reduced-price lunch, and students scoring lower on a baseline version of the NWEA MAP assessment.

### Exhibit ES3. Percentage of Students Failing at Least One Core Course in Ninth Grade



*Source:* AIR calculations from school-provided administrative data.

*Note.* Differences are statistically significant at the  $p < .001$  level for the full sample, female students, male students, and minority students. Differences are statistically significant at the  $p < .05$  level for white students.

### **What Do These Findings Mean?**

This evaluation found that BARR is an effective model. It was implemented with fidelity, and it improved academic outcomes and student experiences. Because we measured BARR's impact at the end of the first year of a 3-year engagement with the study schools, it is possible that these effects may grow larger over time. The evaluation was not designed to detect longer term academic impacts. Also, the BARR model normally includes the entire ninth-grade cohort or the entire school. The within-school design of the evaluation limited the number of teachers and students who participated in the model and may have hampered the implementation and potential impact of some program components.

A 66-school, i3-funded scale-up study is currently under way. That study will assess whether the BARR model can be scaled up effectively and whether the favorable impacts found in this study and in previous research can be sustained in a larger sample of high schools, all of which are selected for their relatively low academic performance relative to the average in their states. Results from that scale-up study will be available in 2021.



## Overview of the BARR Model and Evaluation Design

Building Assets, Reducing Risks (BARR) is a comprehensive, strength-based approach that uses eight interlocking strategies, including the use of real-time data, to build intentional staff-to-staff, staff-to-student, and student-to-student relationships. On the basis of their experiences with the model so far (e.g., see Corsello & Sharma, 2015), the program developers expect that schools that implement the program with fidelity will see notable changes in teacher experiences (e.g., self-efficacy, view of schools and colleagues) as well as student school experiences, behaviors, and, over time, academic outcomes. After completing a successful development grant, the U.S. Department of Education’s Investing in Innovation (i3) program provided the BARR developers with a validation 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 eleven 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 3-year program. The schools participated in the study in three distinct one-year cohorts (three in 2014–15, three in 2015–16, and five in 2016–17).

The impact evaluation uses a multisite, within-school, student-level randomized controlled trial (RCT) to estimate the impact of the BARR program. Individual ninth-grade students in the high schools participating in the evaluation were randomly assigned to the BARR model or to a “business as usual” 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.

This report first provides a brief description of the BARR model, followed by an evaluation design section that describes our measurement plan and characteristics of the schools and students participating in the study. Subsequent sections describe the implementation of BARR by the study schools, differences in the experiences and attitudes of BARR and control teachers, and impacts on student academic and nonacademic outcomes. We conclude with limitations and a summary of findings.

### What Problem Does BARR Address?

The transition from eighth grade to ninth grade is a critical point for students that can set them on a path toward successfully graduating from high school or dropping out. Students beginning high school commonly experience increased stress and behavior problems alongside declines in grades, attendance, interest in school, and perceptions of their academic competence and self-esteem (Alvidrez & Weinstein, 1993; Benner & Graham, 2009; Reyes, Gillock, Kobus, & Sanchez, 2000; Uvaas & McKeivitt, 2013). Research indicates that grade declines, course failure, and

increased absenteeism in ninth grade are strong predictors of dropout in later years (Allensworth & Easton, 2005; Christle, Jolivet, & Nelson, 2007; Easton, Johnson, & Sartain, 2017). And even for students who do not drop out of high school, a positive ninth-grade experience is predictive of later academic success, including passing courses and higher overall GPA, which are predictors of college readiness (Allensworth & Easton, 2005; Bruce, Bridgeland, Fox, & Balfanz, 2011; Dweck, Walton, & Cohen, 2014; Hartman et al., 2011; Norbury et al., 2012).

The developers of BARR designed their model to enable schools and teachers to better support students at a critical point in their academic lives and throughout their high school experience, which the BARR developers expect to translate into lasting effects on student academic outcomes. Such support is centered on enhancing staff-to-staff, staff-to-student, and student-to-student relationships. Strong student-teacher relationships in school require a shared understanding and appreciation of students' strengths, challenges, interests, and circumstances, which is difficult to maintain in the departmentalized structure of a typical large high school (Lieberman, 1990). The changes in the school environment that accompany student transitions into high school are believed to explain much of the increase in academic and behavioral challenges that students experience in ninth grade (Cauley & Jovanovich, 2006; Stoker, Liu, & Arellano, 2017). BARR is designed to mitigate these changes in the school environment and to effectively support student development and learning in both large and smaller high schools.

### **Focus of the BARR Model**

BARR is built upon an extensive body of research that has found that teacher effectiveness is essential to student success and that teacher effectiveness is influenced by a number of intersecting variables, including the extent to which teachers have (a) opportunities for peer-to-peer learning and collaboration; (b) capacity to build positive, intentional relationships with their colleagues and students; and (c) improved communication with families, treating them as active partners along with school leaders (Goddard, Goddard, & Tschannen-Moran, 2007; Henderson & Mapp, 2002). BARR creates and supports teacher teams using several inter-related approaches designed to impact these three critical variables.

**Peer-to-peer learning and collaboration.** The first step in the BARR model is to reorganize the ninth-grade master schedule into distinct cross-subject groups in which students attend their core classes (English language arts [ELA], mathematics, science, and/or social studies) together. The teachers of these core classes share a common planning time, which is intended to create and reinforce opportunities for peer-to-peer learning and collaboration. Such increased collaboration, in turn, is intended to create opportunities to extend the reach of experienced teachers into their fellow teachers' classrooms. Social network analyses have identified significant positive "spillover effects" that occur when teachers work together (Penuel, Sun,

Frank, & Gallagher, 2012); such effects have been directly linked to student academic outcomes. Jackson and Bruegmann (2009) found that teachers with effective colleagues experience peer-related learning that results in improved test scores for their students, and several studies have found a positive relationship between teacher collaboration and student achievement in reading and mathematics (Goddard et al., 2007; Ronfeldt, Farmer, McQueen, & Grissom, 2015).

**Relationship building.** The BARR program provides intensive professional development designed to increase the relationship-building skills of teachers and school administrators in BARR schools. Researchers have noted that restructuring a school environment does not necessarily create positive outcomes unless attention is paid to the nature and quality of teacher interaction (Graham, 2007; Levine, 2010). BARR seeks to address this issue by offering teachers a professional development curriculum built upon positive youth development principles that build the relationships, opportunities, values, and skills for school and life success (Benson, 2007).

**Environmental support.** The BARR program aims to create an environment in which teachers feel supported. A lack of support can cause isolation and emotional stress, which in turn can cause teachers to experience depression, exhaustion, reduced empathy, and a lack of feelings of personal accomplishment (Halbesleben, 2006; Jennings & Greenberg, 2009; Mahan et al., 2010). Teachers who are burned out tend to disengage from the profession, distancing themselves emotionally and professionally from students and colleagues (Mahan et al., 2010). Research shows that teachers who work in supportive environments are more likely to improve their effectiveness over time than those who are in less supportive environments (Johnson, Kraft, & Papay, 2014). BARR seeks to prevent or reverse such burnout by supporting teachers' development of their professional identities and creating healthy working environments in which teachers feel supported by peers and administrators.

### **Development of the BARR Model**

The BARR model was developed at 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 funding from a 2010 i3 Development Grant, BARR was implemented in two schools in Maine and one school in California. This implementation was rigorously evaluated with an RCT in the California school, resulting in statistically significant positive effects on reading achievement, mathematics achievement, and the number of core credits earned (Corsello & Sharma, 2015).

On the basis of these results, BARR developers were awarded an i3 Validation Grant in 2013 to support a rigorous study of the BARR model across a range of geographic regions, including eleven schools in Maine, California, Minnesota, Kentucky, and Texas.<sup>1</sup> Following promising results from the first two cohorts of schools in this validation study, BARR developers applied for and were awarded an i3 scale-up grant in 2017. BARR has grown substantially over the years. It is now serving more than 100 schools in fifteen states and Washington, DC, training more than 2,000 teachers and reaching more than 22,000 students annually.

## **Components of the BARR Model**

The BARR model requires schools to create the structural and organizational conditions necessary to fully integrate student supports into a school’s existing model for addressing nonacademic barriers to learning. Implementation of the BARR model within a school involves restructuring the ninth grade into three- to four-person teacher teams. These teams are then expected to engage in collaborative assessment of all students, problem solving, and planning in weekly block/team meetings. Students identified as high risk in block/team meetings are elevated to risk-review meetings. School staff engage in a collaborative assessment of the progress and challenges of high-risk students and problem solving to address these challenges. These risk-review meetings include additional support staff such as counselors, assistant principals, and school resource officers. As a result, both academic and nonacademic staff have a shared understanding of the assets these students bring to school, the academic and nonacademic risks they face, and the interventions that are under way to support them.

In each school, these meetings and other components of the BARR model are supported by a designated BARR coordinator whose time commitment ranges from 50 to 100 percent full-time equivalent (FTE), depending on the size of the school. Once schools have integrated these structural conditions, school staff and leaders engage in continuous professional development provided by the BARR developers, including annual training and in-situation coaching, phone-based support, quarterly mentoring visits, and technology-enabled learning opportunities. The BARR developers also maintain an ongoing learning community of BARR schools, which come together in annual, in-person program meetings to share their implementation experiences and challenges.<sup>2</sup> Through these ongoing learning opportunities, teachers, administrators, and other school staff learn how to (a) engage the “whole student” in day-to-day interactions; (b) learn how to administer I-Time, a social and emotional development curriculum detailed later; and (c) connect with other caring adults, such as family members and district staff, to support the needs of all students.

---

<sup>1</sup> In total, 46 schools across thirteen states were served as part of the i3 Validation Grant that funded this evaluation. In addition to the eleven schools included in the RCT detailed in this report, 35 additional schools, mostly in rural settings, received support from BARR to implement the model to assess and inform the broad adoptability of BARR.

<sup>2</sup> During the first two years of this evaluation, these meetings occurred semi-annually.

The BARR developers organize the activities described above around the following eight strategies (also referred to as “*components*” throughout this report):

**Strategy 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. For example, if a student disengages in one teacher’s class but is highly engaged in another teacher’s class, the teachers are supposed to jointly identify what strengths this student brings to the class in which he or she is fully engaged. Then, the teacher with the engagement problem is supposed to reference these strengths when discussing the problem with the student or with his or her parent or guardian. Similarly, teachers are supposed to know 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.

**Strategy 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.

**Strategy 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.

**Strategy 4: Create Groups of Students With Common Teachers.** As mentioned above, 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.

**Strategy 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 challenges 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.

**Strategy 6: Conduct Risk-Review Meetings.** Students who persistently fail or exhibit major attendance or behavioral problems 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.

**Strategy 7: Engage Families in Student Learning.** The BARR model seeks to support ongoing interaction with parents 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 parent outreach components. A similar, assets-first approach that is used in block/team and risk-review meetings is applied to conversations with parents.

**Strategy 8: Engage Administrators.** The BARR program requires ongoing commitment from school leadership (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.

Together, these strategies 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). The BARR developers and participating schools expect 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, or acceptance to more selective colleges). A logic model tying these different outcomes together is provided in Appendix A.

## Evaluation Design

BARR was implemented in 46 schools in thirteen states with funding provided by the i3 Validation Grant, reaching approximately 1,797 teachers and 30,943 students. This study focuses on eleven schools serving high-need students. The schools were recruited to be part of an RCT and agreed to randomly assign their entering ninth-grade students to a BARR group or a business-as-usual control group. The additional 35 “dissemination” schools not examined through this study were followed directly by BARR. For information on these schools’ experiences implementing the program, see the callout box titled “Experiences and Outcomes of the BARR Validation Schools Not Included in the Randomized Controlled Trial,” which appears on Page 46 at the end of the student impact section of this report.

### Study Design

This student-level RCT provides causal estimates of program impact on student academic achievement and experiences for three cohorts of ninth-grade students in eleven schools after one year of BARR implementation. Three study schools participated in the first cohort (2014–15 school year), three schools participated in the second cohort (2015–16 school year), and five schools participated in the third cohort (2016–17). Within each school, ninth-grade students were randomly assigned to take at least three of their core academic classes (i.e., ELA, mathematics, science, and/or social studies) with either BARR teachers (treatment group) or non-BARR teachers (control group).<sup>3</sup> Students in the treatment condition were taught by teams of teachers trained in the BARR model, supported by a BARR coordinator, and assisted by BARR coaches. Students in the control condition were taught by teachers operating in business-as-usual conditions in the school. (These teachers are referred to as *control teachers* in the remainder of this report).

The evaluation addressed the following three confirmatory research questions:<sup>4</sup>

1. How did the BARR model impact the incidence of course failure in core subjects (i.e., English, mathematics, science, and social studies)?
2. How did the BARR model impact student test scores on the Northwest Evaluation Association’s Measures of Academic Progress (NWEA MAP) English language arts test?

---

<sup>3</sup> The only students who were excluded from random assignment were special education students in self-contained classrooms.

<sup>4</sup> These confirmatory research questions cover what the program developers considered 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. The other research questions addressed in this report and related impact estimates are considered exploratory. 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.

3. How did the BARR model impact student test scores on the NWEA MAP mathematics test?

These confirmatory research questions were examined with the full study sample for the primary analyses. Subanalyses estimated BARR’s impacts within student subgroups, including those defined by gender, minority status, English learner status, special education status, whether students qualified for free or reduced-price lunch (FRPL), and student performance on a baseline administration of the NWEA achievement tests. We also examined how the impacts varied across the eleven schools.

In addition to the three outcomes described above, we explored the extent to which assignment to the BARR program impacted several other student outcomes, including GPA, the degree to which students met NWEA growth targets between the pre- and post-tests, their experiences in school, and their behaviors. These exploratory analyses are informed by the following research questions:

4. How did the BARR model impact students’ GPA at the end of ninth grade?
5. How did the BARR model impact the percentage of students meeting NWEA growth targets in reading and mathematics, as measured with the NWEA baseline and follow-up tests administered in the evaluation?
6. How did the BARR model impact the classroom and school experiences of students, including their engagement, their grit, their sense of belonging, and the perceived quality of their relationships with their teachers and peers (i.e., social and emotional learning, supportive relationships, and expectations and rigor)?
7. How did the BARR model impact student attendance (as measured by the incidence of chronic absenteeism)?
8. How did the BARR model impact the likelihood of students being suspended for disciplinary reasons?
9. How did the BARR model impact the likelihood that students would enroll in 10th grade in the same school in the school year following ninth grade?

During the same end-of-ninth-grade time frame when we measured student experiences with a student survey, we also administered a survey to their core subject teachers. These surveys enabled us to answer the following exploratory research questions about the teachers’ experience with BARR.

10. How did BARR teachers and control teachers view their colleagues and the quality of collaborations with them?
11. To what extent did BARR teachers and control teachers use data to inform their instruction?
12. How did BARR teachers and control teachers view their students’ behavior?
13. How did BARR teachers and control teachers view their interactions with parents?



14. How did self-reported self-efficacy differ between BARR teachers and control teachers?

Two final research questions focused on the implementation of BARR in treatment schools. Using the eight BARR strategies introduced above as an organizing framework, we systematically collected program implementation data in each of the eleven study schools to answer the following implementation questions:

15. To what extent were the eight BARR strategies implemented as intended?

16. What were the facilitators and barriers to successful implementation?

## Measurement and Data Collection

We used a variety of data sources to answer the research questions regarding the impact and implementation of the BARR program. Details and instruments are presented in appendices referenced throughout this section.

### Student Outcomes

**Reading and Mathematics Achievement.** At the beginning of the school year, school staff administered two standardized tests to all ninth-grade students: the NWEA MAP assessment in mathematics and reading. These assessment data served as a baseline measure of academic skill for students in the study. At the end of the school year, school staff again administered the two NWEA MAP assessments to all ninth-grade students, and these data provided two of the three measures of academic achievement for the confirmatory impact analyses.<sup>5</sup>

**Course Failure, GPA and Behavior.** In addition to using standardized assessments, we collected administrative data to examine course failure, GPA, and student behavior. Student-level academic outcome data that were collected included core credits earned and grades for all core courses, which enabled us to examine course failure (earning less than a *D* on any core course) and to calculate students' GPAs at the end of ninth grade. The behavioral outcome data collected included attendance records, disciplinary referrals and actions, suspensions, and enrollment in 10th grade.<sup>6</sup> We also collected demographic data for each student, including

---

<sup>5</sup> As this study progressed from the first NWEA MAP assessment for Cohort 1 in fall 2014 to the last assessment for Cohort 3 in spring 2017, the NWEA MAP assessments became less relevant to participating schools, causing problems with the administration of the follow-up assessments in six of the eleven schools in the evaluation. These problems included low participation rates, differences in participation between the BARR and control groups, and students not taking the assessment seriously, as evidenced by many spending too little time to obtain a valid score. In some schools, the NWEA tests were administered at the very end of the school year, after all high-stakes testing and exams were completed. In one school, the administration was so compromised that we excluded that school from the NWEA impact analyses altogether. To assess the sensitivity of the NWEA impact analyses to student attrition, we conducted sensitivity analyses that included only five schools in which more than 65 percent of the students participated in the follow-up assessment; the difference in participation between the research groups was less than 5 percent. We found that the impact estimates from these sensitivity analyses led to the same conclusions as those conducted with the larger sample of ten schools. Details are provided in Appendix G.

<sup>6</sup> Some of the data on disciplinary actions and suspensions were not available from all study schools. These limitations are detailed in the section of this report in which we discuss BARR's impact on these outcomes.

race/ethnicity, gender, socioeconomic status, eligibility for special education services, and limited English proficiency status.

**Student Experiences.** We administered student surveys in the spring semester each year to assess the effects of BARR on student experiences in school. BARR and control group students received and completed the same survey instrument (Appendix B). We used the survey responses to create scale scores for the following six different dimensions:<sup>7</sup>

- **Student Expectations and Rigor.** Students believe teachers have high expectations for their performance, provide clear guidelines, and encourage them to be successful.
- **Student Engagement.** Students actively prepare for and participate in class activities, ask questions, and are interested in the lesson.
- **Supportive Relationships.** Students feel that their teachers are supportive of their interests and invested in their emotional well-being.
- **Social and Emotional Learning.** Students believe their classmates are capable of working through disagreements, managing their emotions, and acting responsibly.
- **Sense of Belonging.** Students feel respected, accepted, and understood by their peers and feel that they belong with their classmates.
- **Grit.** Students are able to focus on and work toward meeting goals, even when they encounter setbacks.

### ***Student Sample and Attrition***

Exhibit 1 summarizes the samples of BARR and control students in the originally randomized sample and the number of those students for whom we successfully collected outcome data. For the NWEA outcomes, this table is limited to ten schools because follow-up NWEA administration was unsuccessful in one school. The table also shows what percentage of the original sample these students represented. Student attrition (the inverse of this percentage) ranged from 33.7 percent for the student surveys to 18.8 percent for the course failure data. BARR students were more likely to answer the student survey than control students. Sample attrition for these outcomes (relative to the originally assigned sample) was 28.7 percent for the BARR group and 37.4 percent for the control group. See Appendix C for the sample sizes and attrition rates for each individual student confirmatory outcome measure.

---

<sup>7</sup> To create the scores, we used a scaling process known as the *Rasch model* for ordered response categories (Andrich, 1978; Rasch, 1980; Wright & Masters, 1982). As part of this process, scores were standardized to have a mean of 50 and a standard deviation of 10.

**Exhibit 1. Analytic Sample by Data Source and Group**

Outcome	BARR			Control			Total		
	Assigned Sample (n)	Analytic Sample (n)	%	Assigned Sample (n)	Analytic Sample (n)	%	Assigned Sample (n)	Analytic Sample (n)	%
NWEA Reading (10 schools)	1,667	1,147	68.8	2,271	1,448	63.8	3,938	2,595	65.9
NWEA Mathematics (10 schools)	1667	1,173	70.4	2,271	1,542	67.9	3,938	2,715	68.9
Course Failure (11 schools)	1,785	1,467	82.2	2,383	1,916	80.4	4,168	3,383	81.2
Student Experiences (11 schools)	1,785	1,273	71.3	2383	1,491	62.6	4,168	2,764	66.3

Source: AIR calculations from NWEA scores, school-provided administrative data, and AIR-administered surveys.

**Teacher Experience Outcomes**

To capture the experiences of teachers serving students in the BARR and control groups, we administered a survey at the end of the spring semester to core subject teachers who taught ninth-grade students during the school year. Teachers of BARR and control group students received and completed the same survey instrument (Appendix B). The following constructs were measured:

- **School Supports.** Teachers feel the school provides students and staff with necessary resources and support services.
- **Self-Efficacy.** Teachers believe they have the ability and the 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.
- **Data Use.** Teachers have access to data, discuss data, and use data to differentiate instruction for students.
- **View of Students’ Observed Behavior, Commitment, and Attitudes.** What teachers observe students doing in classrooms related to goal setting, motivation, and completing work.
- **Perception of Students’ Behavior, Commitment, and Attitudes.** What teachers think students would do to help their peers or make the right choices in a given situation.
- **Interaction With Parents.** Teachers feel confident engaging with parents and helping them understand what students need to learn.

- **Student Accountability.** Teachers believe students are accountable for their own learning and grades and should meet all deadlines.

### **Teacher Sample and Attrition**

Forty-eight BARR teachers and 65 control teachers completed the survey, for a total response rate of 64.9 percent (84.2 percent for BARR teachers and 55.6 percent for control teachers). The response rate for Cohorts 1, 2, and 3 BARR teachers ranged between 79.2 percent and 92.9 percent and the response rate for each of the three cohorts of control teachers ranged from 47.6 percent to 60.4 percent. Note that these teachers were not randomized to a BARR group or a control group, which would have been very difficult logistically and in many cases would have been impossible, considering work assignment and licensure rules. Therefore, while BARR participation may have caused differences in outcomes between teachers in these two groups, the lack of randomization means that we cannot conclusively attribute all observed differences to BARR.<sup>8</sup>

### **Program Implementation**

We used two data sources to measure the fidelity of BARR implementation: ratings from the BARR coordinator structural component review interview and site-visit observations of three BARR activities (block/team meetings, I-Time lessons, and risk-review meetings). Both activities took place during annual spring site visits for each cohort (2015–17). During the structural component review interview with AIR researchers, the BARR coordinator shared and explained his or her ratings for the school on a number of relevant indicators<sup>9</sup> for each of the eight key components of the BARR model. In addition, two evaluators directly observed and rated block/team meetings, I-Time lessons, and risk-review meetings using fidelity rubrics that were first developed for the 2010 i3 development project and subsequently revised for this study.

For each of the evaluated implementation components, the research team worked with the BARR developers to establish thresholds defining the level at which the developers considered the BARR model to be adequately implemented at the school. For each component, we assigned a score of 0 to a school if the school failed to reach the minimum rating level and a score of 1 if the school reached the minimum rating level. Averaging component scores yielded an overall measure (index) of fidelity for each school, for each BARR strategy, and for the program as a whole.

In addition to these ratings, we conducted interviews with school administrators and core teachers in each school during annual spring site visits (2015–17).<sup>10</sup> Interview questions focused on the eight BARR strategies for BARR teachers as well as on the use of similar strategies by

---

<sup>8</sup> The i3 scale-up grant currently under way includes an evaluation with school-level randomization, which will enable us to obtain fully experimental estimates of the impacts of BARR on teacher and school-level outcomes.

<sup>9</sup> Indicators are measures or values that provide information regarding the presence or state of a program component within a school.

<sup>10</sup> We did not interview all core teachers in each school but a convenience sample of 106, approximately evenly distributed across schools and subjects.

non-BARR teachers. Teachers were also asked whether there had been any school-level challenges that year, any challenges specific to BARR, or, in the case of the BARR teachers, any additional supports they would like for future implementation. These data were used to understand the facilitators and barriers to successful implementation.

### Baseline Characteristics of the Study Schools and Their Students

The eleven study schools included two rural schools in Maine, one rural school in Kentucky, six suburban schools in California, one suburban school in Minnesota, and one urban school in Texas. Overall, 75 percent of students were minorities, 79 percent were FRPL eligible, 30 percent were English learners, and 8 percent had special education status.

Exhibit 2 shows that there was considerable variation in the demographic background characteristics of the students across the study schools. The six California schools and the one Texas school included much greater percentages of students of color, English language learners, and students from low-income families. Students in the two schools in rural Maine and the school in Kentucky were predominantly White, and about half or fewer of them qualified for FRPL. The study school in Minnesota had the lowest rate of students from low-income families, with fewer than one in five students eligible for FRPL.

**Exhibit 2. Characteristics of Schools in Cohorts 1, 2, and 3**

Cohort	School	State	Locale	Students of Color (%)	English Language Learners (%)	Special Education Students (%)	Eligible for Free or Reduced-Price Lunch (%)
Cohort 1	School A <sup>a</sup>	California	Suburban	94.0	15.4	11.2 <sup>b</sup>	89.0
	School B <sup>c</sup>	Maine	Rural	7.0	< 1	17.7 <sup>d</sup>	39.2
	School C <sup>a</sup>	California	Suburban	72.4	8.3	16.6	80.7
Cohort 2	School D <sup>a</sup>	Maine	Rural	3.5	0.3	12.5 <sup>d</sup>	53.4
	School E <sup>e</sup>	Minnesota	Suburban	25.0	2.6	9.6	19.7
	School F <sup>f</sup>	California	Suburban	93.1	14.3	10.7	80.6
Cohort 3	School G <sup>g</sup>	Texas	Urban	71.5	18.5	08.6	59.0
	School H <sup>h</sup>	California	Suburban	93.8	13.4	11.2	83.5
	School I <sup>i</sup>	Kentucky	Rural	1.9	NA	NA	36.2
	School J <sup>h</sup>	California	Suburban	95.4	7.8	7.6	75.3
	School K <sup>h</sup>	California	Suburban	89.3	10.1	10.7	73.1

Sources: <sup>a</sup>California Department of Education Educational Demographics Unit Database (2014–15); <sup>b</sup>California School Accountability Report Card (2013–14); <sup>c</sup>Maine School Accountability Report Card (2014–15); <sup>d</sup>Maine Department of Education Data Warehouse (2014–15); <sup>e</sup>Minnesota Report Card (2015–16); <sup>f</sup>Ed-Data Education Data Partnership (2014–15); <sup>g</sup>Texas School Accountability Report Card (2015–16); <sup>h</sup>California School Accountability Report Card (2015–16); <sup>i</sup>Common Core of Data (2015–16).

After randomization, we determined that the two research groups in the study sample were equivalent in gender, ethnicity, ELL status, special education status, FRPL status, and NWEA pretests (mathematics and reading).<sup>11</sup> We evaluated baseline equivalence separately for each analytic sample, as detailed in Appendix D. The results indicated that the BARR group and the control group were closely balanced on pre-intervention background characteristics for all analytical samples.

The combined sample for Cohorts 1, 2, and 3 included 4,168 ninth-grade students: 1,785 BARR students and 2,383 control students (Exhibit 3). There were more control students than BARR students in the study sample because capacity constraints in large high schools (e.g., School F, School G, and School K) were addressed by applying a 1:2 or 1:3 BARR/control random-assignment ratio.<sup>12</sup>

**Exhibit 3. Randomized Student Sample for Cohorts 1, 2, and 3**

Cohort	Assigned Sample	BARR	Control	Total
Cohort 1	School A	321	321	642
	School B	104	104	208
	School C	179	178	357
Cohort 2	School D	82	81	163
	School E	83	84	167
	School F	211	422	633
Cohort 3	School G	150	381	531
	School H	140	185	325
	School I	118	112	230
	School J	200	162	362
	School K	197	353	550
<b>Cohort Total</b>		<b>1,785</b>	<b>2,383</b>	<b>4,168</b>

Source: AIR sample from school-provided administrative data.

<sup>11</sup> The difference between the BARR group and control group for each baseline characteristic was less than 0.25 standard deviations, thus meeting the What Works Clearinghouse (WWC) baseline equivalence guidelines.

<sup>12</sup> This variation in the random-assignment ratio across schools caused the BARR treatment status to be correlated with school in the full study sample. We used statistical controls to account for this correlation in our impact analyses. Anyone using future public-use data from this study should refrain from making unadjusted BARR-control comparisons that do not account for this variation, because such comparisons will result in misleading findings.

## Did Study Schools Implement BARR as Intended?

Schools participating in BARR are offered 3 years of professional development and coaching. This evaluation focuses on the implementation and impacts of BARR during the first of these 3 years. The eleven study schools were offered three 2-day, on-site BARR training sessions at each school for all educators involved in the BARR model. These 2-day trainings were scheduled once each year, for a total of 6 days of on-site training across the 3 years. The first year of professional development focused on understanding the BARR model, adopting a whole student approach, identifying and leveraging student strengths, and practice facilitating the I-Time lessons. Alongside on-site training, BARR schools received BARR implementation guides with a DVD/CD, the I-Time classroom curriculum, and a video training program to assist schools with implementing block/team meetings, risk review, and I-Time lessons.

The BARR educator who conducted the training at each school served as the school's dedicated BARR coach. As stated previously, BARR coaches provided in-situation coaching, phone-based support, quarterly mentoring visits, and technology-enabled learning opportunities. Each school was expected to participate in three on-site coaching visits in the first year, to be followed by two on-site visits in the second year and one on-site visit in the third year. (The visits beyond Year 1 fell outside the scope of this evaluation). During these visits, the coach would observe BARR implementation and would provide feedback on how the school could increase or maintain fidelity to the BARR model. In addition to these scheduled coaching activities, each school had unlimited access to the BARR coach through virtual coaching.

In addition to receiving coaching support, BARR coordinators were asked to participate in monthly professional learning community webinars with other schools across the country implementing the BARR model to share best practices, challenges, and strategies. Schools also were strongly encouraged to send participants (BARR coordinators and teachers) to semiannual, in-person program meetings to share their implementation experiences and challenges and to interact in person with other BARR schools and with BARR program staff.

In addition to these BARR-administered program activities, the schools were expected to staff a BARR coordinator (at 50 to 100 percent FTE depending on the school size) and were expected to free up teacher and school administrator time for block/team meetings and risk-review meetings. These meetings were part of teachers' regular planning time, or teachers were paid for their attendance at meetings after school. Most importantly, BARR schools were expected to ensure their schedule enabled core teachers to share at least 80 percent of the same ninth-grade students.

To measure implementation fidelity, we documented program activities in each of the eleven study schools, organized around the eight BARR strategies/components described earlier. We

averaged the BARR coordinator and observer indicator ratings to create a composite score for three components: I-Time, block/team meetings, and risk-review meetings. For the remaining five components, the average BARR coordinator indicator rating was used. For each school, this component score was compared against a predetermined threshold for adequate implementation. If the 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. These thresholds were determined on the basis of theory and practice at the beginning of the study, in consultation with the program developers. Appendix E provides information on these thresholds and greater detail on how we calculated fidelity scores. Across the different schools, we subsequently established an overall “yes/no” indicator for implementation with fidelity in each cohort. This indicator was based on at least two-thirds (67 percent) of schools being rated as adequate implementers for that year.

### **To What Extent Were the Eight BARR Strategies Implemented as Intended?**

All BARR key components were adequately implemented in at least two-thirds of the schools (Exhibit 4). Therefore, we conclude that the BARR model was adequately implemented in this study. Across the three cohorts, we found that all study schools (100 percent) were able to restructure ninth grade, focus on the whole student, and engage families with adequate fidelity. All schools but one (91 percent) were able to implement block/team meetings, risk-review meetings, and contextual supports with adequate fidelity. Similarly, all but two schools (82 percent) were able to implement I-Time with adequate fidelity. Professional development was the most challenging, with three schools not meeting school-level fidelity. These three schools were rated as *not yet* or *emerging* for holding regular monthly meetings on implementation and trainings on how to address student needs. The remaining eight schools met school-level fidelity (73 percent) as a result, the program met adequate overall fidelity for this component.

### **Implications of the Implementation Findings and Study Design for the Treatment Contrast**

The generally high levels of implementation fidelity in this study mean that, at least from the perspective of the BARR treatment group after 1 year of implementation, this evaluation provides a fair assessment of the impact of the BARR model. There are, however, some additional concerns about the validity of the treatment contrast due to the design and implementation of the within-school RCT design:

1. All eleven study schools implemented BARR for the first time during the study year, which means that the program was not yet in a well-established, “steady state” operating environment in these schools during the year in which it was evaluated.



**Exhibit 4. Reporting Fidelity for Each BARR Key Component: All Three Cohorts**

BARR Key Component/ Strategies on Logic Model	Definitions		Fidelity Findings							
			Cohort 1 2014–15 (3 Schools)		Cohort 2 2015–16 (3 Schools)		Cohort 3 2016–17 (5 Schools)		All Three Cohorts 2014–17 (11 Schools)	
	Definition of Adequate Implementation	Definition of Implementation With Fidelity at Program Level	Schools With Adequate Implementation (%)	Implemented With Fidelity (Yes/No)	Schools With Adequate Implementation (%)	Implemented With Fidelity (Yes/No)	Schools With Adequate Implementation (%)	Implemented With Fidelity (Yes/No)	Schools With Adequate Implementation (%)	Implemented With Fidelity (Yes/No)
Professional Development	Mean of three indicators (4.7 or higher)	Two-thirds (67%) of schools rated as <i>adequate</i>	67	Yes	100	Yes	60	No	73	Yes
Restructuring Ninth Grade	Mean of five indicators (5.3 or higher)	Two-thirds (67%) of schools rated as <i>adequate</i>	100	Yes	100	Yes	100	Yes	100	Yes
Whole- Student Emphasis	Mean of four indicators (5.5 or higher)	Two-thirds (67%) of schools rated as <i>adequate</i>	100	Yes	100	Yes	100	Yes	100	Yes
Block/Team Meetings	Mean of seven indicators and observations (5.6 or higher)	Two-thirds (67%) of schools rated as <i>adequate</i>	100	Yes	100	Yes	80	Yes	91	Yes

BARR Key Component/ Strategies on Logic Model	Definitions		Fidelity Findings							
			Cohort 1 2014–15 (3 Schools)		Cohort 2 2015–16 (3 Schools)		Cohort 3 2016–17 (5 Schools)		All Three Cohorts 2014–17 (11 Schools)	
	Definition of Adequate Implementation	Definition of Implementation With Fidelity at Program Level	Schools With Adequate Implementation (%)	Implemented With Fidelity (Yes/No)	Schools With Adequate Implementation (%)	Implemented With Fidelity (Yes/No)	Schools With Adequate Implementation (%)	Implemented With Fidelity (Yes/No)	Schools With Adequate Implementation (%)	Implemented With Fidelity (Yes/No)
I-Time	Mean of six indicators and observations (5.3 or higher)	Two-thirds (67%) of schools rated as <i>adequate</i>	100	Yes	67	Yes	80	Yes	82	Yes
Risk Review	Mean of six indicators and observations (4.8 or higher)	Two-thirds (67%) of schools rated as <i>adequate</i>	100	Yes	100	Yes	80	Yes	91	Yes
Contextual Support	Mean of seven indicators (4.0 or higher)	Two-thirds (67%) of schools rated as <i>adequate</i>	100	Yes	100	Yes	80	Yes	91	Yes
Parent Involvement	Mean of six indicators (3.3 or higher)	Two-thirds (67%) of schools rated as <i>adequate</i>	100	Yes	100	Yes	100	Yes	100	Yes

2. Because the program was limited to half or fewer of the ninth-grade students and core subject teachers, the implementation in the study schools lacked some of the program experiences and activities that are available when BARR is implemented universally. These include, for example, special ceremonies and awards, awareness campaigns, and similar schoolwide activities. Moreover, partial implementation of BARR and the need to maintain a within-school treatment contrast may have reduced the program’s potential to impact teacher-to-teacher relationships (e.g., via department-level meetings or informal gatherings) as well as limited its impact on the larger school climate.
3. In each of the schools, some of the staff included in BARR professional development and in other activities like block/team meetings and risk-review meetings interacted with students in both research groups (treatment and control). Such staff might include assistant principals, counselors, reading specialists, and school resource officers. Although these staff agreed to limit specific BARR activities to the treatment group, it would have been difficult for them not to use some of their BARR-related skills and approaches in their interactions with the control group as well. Therefore, some contamination of the “business as usual” control group experience was likely.

As a result, it is likely that the impact estimates from this evaluation are attenuated relative to the hypothetical effects of a full, 3-year, schoolwide implementation of the BARR model. The ongoing i3 scale-up evaluation of BARR will partially address this issue through school-level random assignment that maintains a treatment contrast for the first year of implementation.

### **What Were the Facilitators of and Barriers to Successful Implementation?**

We conducted supplemental qualitative interviews to further contextualize the implementation experience. The BARR coordinators, teachers, and school administrators were asked questions about implementing the eight BARR strategies and what they identified as the primary facilitators for and barriers to implementing the model. Overall, these interviews revealed that most BARR teachers viewed their overall experience with BARR positively. For example, one teacher shared the following:

“If you were to talk about successes, we would be here for another hour. There [are] always challenges, but we’ve always gotten through it. Next year, it’s going to be a lot better if this year went great. If [in] this year we have great success stories, I’m just so excited for next year.”

BARR teachers also reported receiving positive feedback from students and parents. One BARR teacher shared the following:

“Parents are saying, ‘I feel like you care about my child, that you’re on top of things, that you’re communicating with us and with other teachers.’ The kids ask almost daily [if we are] going to do an I-Time. They love I-Times. I think they are generally happy.”

BARR teachers in most schools stated that they benefited from the relationships that were built with students, between students, and with other teachers outside their own department and could give clear examples of how that relationship building had a positive impact on their classroom practice. In most schools the BARR team, beyond the teachers, included other professionals with strong investment in the work of the team, such as a school counselor, who learned many details of students’ lives through her role in the school, which provided additional information that was useful to the BARR team. Such cross-function collaboration was a facilitator of the successful implementation in those schools.

Almost all interviewed BARR teachers identified the support of the BARR coordinator as the primary facilitator of successful implementation. Teachers saw the person in this role as a leader who provided direct support to teachers, organized and facilitated BARR activities, maintained frequent communication, served as the data coordinator, and helped keep teachers on track and motivated. Most teachers indicated that implementing BARR would have been much harder if they had not had a BARR coordinator to facilitate the process.

Another facilitator of program implementation that many teachers mentioned was the increased sense of community that BARR created in their school. Teachers reported having stronger relationships with their students and colleagues than they had in previous years (before their school’s adoption of BARR). One BARR teacher expressed the following sentiment:

“I just enjoyed my freshmen this year. I was really close with them. I feel like they were close with one another. It seemed like a family atmosphere.”

Another BARR teacher said that students seemed to recognize the attention teachers gave to their academic success and that the students felt more comfortable approaching teachers for help. Such reciprocal, positive feedback and engagement helped teachers and administrators remain committed to the implementation of BARR in their classrooms and schools.

Some BARR teachers, BARR coordinators, and school administrators identified barriers to successful implementation in the first year that required some adjustment by the school staff. First, teachers mentioned the amount of time needed for block/team meetings, I-Time sessions, and other program-related tasks and activities as a significant barrier, especially initially. For example, BARR teachers in one school expressed concerns about running out of time to cover their regular curriculum because of I-Time. They also mentioned how the BARR block/team

meetings represented time "cut out of other places," such as their preparation time. These teachers reported having to do more work at home as a result.

Lack of time and focused attention also sometimes limited self-reflection about BARR implementation and related professional development needs. For example, one BARR coordinator reported that the BARR teachers in her school had not committed the necessary time for monthly assessment of their BARR model practices, nor had they identified areas of needed professional development.

Aside from time pressures and related constraints, other barriers identified by school staff included understanding ways to use data to track student progress, coordinating scheduling and student grouping, navigating limited parent engagement, and dealing with the learning curve associated with a new program.

### **How Could Implementation Be Improved?**

During interviews, BARR teachers identified additional supports that would be helpful for continued implementation. First, teachers highlighted the importance of improving the data collection and tracking processes. One teacher expressed the following:

“My fingers are crossed that BARR continues . . . [to] create better software . . . [to] run reports on real-time student data rather than having teachers fill in a lot of this data.”

Second, aside from providing more time for BARR activities, some teachers expressed a need to encourage more focused attention by all participants during meetings to streamline consensus decision making. A mathematics teacher offered the following explanation:

“The more that the administration and the grant personnel can budget to help teachers spend more time intervening with kids, the better the BARR program will be.”

Lastly, teachers reported that having more opportunities for training and modeling of the program (e.g., more time and training to prepare for I-Time lessons) would improve implementation.

## Did BARR Change Teacher Experiences and Teacher Practice?

An objective of the BARR model is to change how teachers view and interact with their students and with each other. By creating structures and activities to bring teachers together and to deepen teachers' relationships with their students, BARR aims to enhance teacher efficacy and student engagement. We collected survey data to examine differences in the experiences of the BARR teachers and control teachers. These surveys were introduced earlier and are available in Appendix B.

We surveyed a total of 113 ninth-grade teachers in the eleven study schools (48 BARR teachers and 65 control teachers). Fifty-eight percent of these teachers were female, 37 percent were male, and five percent chose not to report their gender. Instructional experience was well distributed; 25 percent of teachers had 1 to 3 years of teaching experience, 28 percent had 4 to 10 years of teaching experience, and 42 percent had eleven or more years of teaching experience.

Exhibit 5 presents the results from a statistical comparison between the responses of the BARR and control teachers on the eight teacher survey measures introduced in the evaluation design. The table shows statistically significant differences between BARR and control teachers for six of the eight survey measures. BARR teachers had more positive views about their colleagues and collaborating with them (effect size [ES] = 1.42), and they reported greater levels of data use to inform their instruction (ES = 1.14). BARR teachers were also more positive in their view of school supports (ES = 0.61), and they reported greater self-efficacy (ES = 0.51) than their control group counterparts. BARR teachers perceived student behavior more positively (ES = 0.48) and reported more positive observed behavior from students (ES = 0.46).

### How to Read Tables Comparing BARR and Control Group Outcomes

Exhibit 5 (on the next page) and many subsequent exhibits in this report include seven columns, which together contain all the information necessary to describe and interpret differences between the BARR group and the control group on a range of outcomes. The **first column** describes the outcome being compared or the subgroup of students for which an outcome is shown. The **second column** (labeled *N*) displays the sample size, which varies depending on the data source, the subgroup definition, and any missing data affecting the outcome. The **third and fourth columns** display the average outcome levels for the BARR group and the control group. The **fifth column** describes the difference between those two, which is the impact of BARR in tables covering student outcomes. (In Exhibit 5, which shows teacher-level outcomes, this difference is not an estimate of the impact of BARR, because teachers were not randomly assigned to BARR or control status). The numbers in the fifth column have asterisks, daggers, and double daggers attached to them if the difference between BARR and control is *statistically significant*, meaning that the likelihood that the apparent difference is the result of chance is less than 5 percent. The **sixth column** displays a *p*-value, which is the likelihood that the difference is zero. The **seventh and final column** displays a standardized effect size, which is equivalent to the difference divided by the standard deviation of the outcome for the full sample. These effect sizes are comparable across outcomes, even if those outcomes have different scales or dimensions.

**Exhibit 5. Differences in Teacher Experiences in Participating Schools**

Outcomes	N	BARR	Control	Difference	P-Value	Effect Size
Teacher collaboration with and view of colleagues	110	56.68	45.02	11.67‡	0.000	1.42
Teacher use of data	109	55.70	45.68	10.03‡	0.000	1.14
View of the school’s supports	111	53.47	47.55	5.92†	0.002	0.61
Teacher self-efficacy	112	52.92	47.89	5.03†	0.006	0.51
Perception of students’ behavior	111	52.72	48.00	4.72*	0.012	0.48
View of students’ observed behavior	113	52.59	48.09	4.49*	0.016	0.46
Interaction with parents	111	52.07	48.54	3.53	0.058	0.35
View of student accountability	110	50.26	49.80	0.46	0.804	0.05

Source: AIR calculations from AIR-administered teacher surveys.

Note. \* = statistically significant at the  $p < .05$  level; † = statistically significant at the  $p < .01$  level; ‡ = statistically significant at the  $p < .001$  level.

In addition to the survey data presented in Exhibit 5, we collected interview data from a subsample of 106 teachers about changes in their practice, collaboration with other teachers and school staff, and the overall organization of the school. The specific interview prompts are included in Appendix B. These interviews confirmed that BARR teachers perceived moderate or major changes in their teaching practices during the year in which BARR was implemented in their school. For example, all but one BARR teacher interviewed (98.1 percent) reported that using social and emotional learning or youth development activities/lessons changed their teaching practice to a moderate or major extent.<sup>13</sup> Only 45.7 percent of teachers in the control group reported a similar change in their teaching practice during the year. Similarly, BARR teachers were much more likely to report an increased reliance on team meetings (96.4 percent versus 60.0 percent in the control group) and use of the current student referral process (73.1 percent versus 35.0 percent in the control group). Although these interview data underscore the penetration of the BARR model into teacher practice in the study schools, they also indicate that control group teachers sometimes experienced meaningful changes in their teaching practice during the same year as well.

<sup>13</sup> The wording of the interview questions and the response options given to teachers was slightly modified after we completed Cohort 1 interviews. We analyzed Cohort 1 interview data separately from Cohorts 2 and 3 interview data, and the same pattern emerged.

## In Their Own Words: How Teachers Experienced the Impact of BARR on Their Teaching Practice

“I feel like a different teacher. I feel like my relationships with students are different. I feel like I’ve been able to create different connections with students than I have [in] the past.”

“Scheduling . . . was probably a challenge for the school. . . . Planning here is always a little tight [anyway] but getting all the teachers that need to have the third period planning—I think, from what I’ve heard—was a little bit of an issue. But again, I think it’s one of the things that’s helped us be successful, too.”

“Above the meetings and the lessons and things, . . . it’s kind of been empowering to ask kids questions and have conversations with kids. I can think of a handful of kids, maybe four or five students, that I’ve learned things about through iTimes. I’ve learned things through conversations with teachers. Those are students I have historically not had relationships with. That, to me, has been really a good thing. I think I’ve been able to reach kids academically and personally through the structure of BARR that I just wouldn’t have [otherwise]. As a teacher, that’s exciting.”

“I think the cool thing is . . . that not only did it foster really good relationships between students and adults in this school, but it also fostered really great relationships among the teachers . . . and I think that’s been huge because we feel a lot more connected and united, and so we also feel more like we’re a front together, doing this all together.”

“[I]t’s challenging losing prep time . . . so it went from five hours of prep a week to three hours a week. . . . In terms of in the classroom, the I-Time activities would—we would lose about a period a month, about an hour or two a month—because [of] I-Time activities. So that was a little challenging, too, knowing how to adjust the schedule and how to kind of fit in the I-Time.”



## Did BARR Impact Student Achievement, Experience, and Behaviors?

This section describes the impact of BARR after 1 year of implementation, on student academic outcomes, experiences, and behavior. As detailed earlier, the primary confirmatory impact analysis includes three measures of student achievement: whether students failed any core courses in ninth grade and end-of-year NWEA reading and mathematics assessment scores. The exploratory impact of BARR on student experiences included six student survey constructs: supportive relationships, expectations and rigor, engagement, sense of belonging, social and emotional learning, and grit. The exploratory impact of BARR on student behavior includes chronic absenteeism, suspensions, and persistence to 10th grade.

### Impacts on Student Achievement

To determine the impact of BARR on course failure, we collected transcript data for all ninth-grade students in the study schools. We used these data to determine whether students failed any core courses (ELA, mathematics, science, and social studies). The other two measures of academic achievement were based on the NWEA MAP assessments administered by each school at the end of the spring semester. Each student's RIT (Rasch Unit) score on these assessments represents performance across a series of subtests within the content areas of mathematics and reading. As discussed earlier, we present NWEA impacts for ten of the eleven schools in the study.

#### Impact Estimation Methods

We used ordinary least squares regression analysis (for continuous outcomes) and *logit* maximum likelihood analysis (for discrete outcomes) to estimate the impact of assignment to BARR on all student outcomes presented in this report. These analyses compare the outcomes for students in the BARR group to students in the control group, controlling for students' background characteristics at baseline, school of enrollment, and baseline NWEA scores. See Appendix F for more details.

### Did BARR Impact Course Failure?

A significantly smaller proportion of BARR students failed one or more core courses in any semester during their ninth-grade year than their counterparts in the control group. Exhibit 6 shows that, among students in the control group, 40.6 percent failed at least one course (i.e., they recorded an *F* on their report card). BARR reduced that percentage to 29.2 percent, a statistically significant impact of 11.5 percentage points (ES of 0.31). Course failure in ninth grade is an important barrier to long-term student success in high school, because students who fail courses in ninth grade have trouble meeting the requirements to advance to the next grade and may instead continue to fall further behind (Bridgeland, Dilulio, & Morison, 2006).

**Exhibit 6. Impacts on Students Failing One or More Core Courses (Full Sample and Subgroups)**

Outcome/Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>3,383</b>	<b>29.2</b>	<b>40.6</b>	<b>-11.5‡</b>	<b>0.000</b>	<b>0.31</b>
Female	1,643	24.1	31.0	-7.0‡	0.001	0.21
Male	1,740	34.2	49.5	-15.3‡	0.000	0.38
Minority	2,497	32.1	45.8	-13.7‡	0.000	0.35
White	886	20.7	25.8	-5.1*	0.041	0.17
Free or Reduced-Price Lunch	2,640	30.8	43.8	-13.0‡	0.000	0.34
Not Free or Reduced-Price Lunch	722	20.4	27.5	-7.1*	0.013	0.24
English Learners	1,071	31.9	43.2	-11.3‡	0.000	0.29
Non-English Learners	2,312	27.8	39.7	-11.9‡	0.000	0.32
Special Education	264	47.7	54.4	-6.7	0.227	0.16
Non-Special Education	3,119	27.6	39.5	-11.9‡	0.000	0.33

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.

Subgroup analyses showed statistically significant impacts on this outcome for most subgroups of students. Impacts were strongest for male students (ES = 0.38), students of color (ES = 0.35), and students eligible for FRPL (ES = 0.34). All three of these groups experienced higher than average rates of course failure in the control group. As a result, Exhibit 6 shows evidence of a significant reduction in the achievement gap along these dimensions, especially between White students and students of color. In the control group, students of color were 20 percentage points more likely to fail a core course than White students. In the BARR group, this gap shrank to 11.4 percentage points, which means that BARR eliminated almost half of this achievement gap for this outcome.

In contrast, we found no meaningful differences in BARR’s impacts on course failure across students’ English learner status (an ES of 0.29 for English learners and an ES of 0.32 for non-English learners). For one pair of subgroups, special education students versus non-special education students, the observed gap was larger for BARR students (an ES of 0.16 for special education students and an ES of 0.33 for non-special education students). The bottom line, however, is that all subgroups appeared to benefit from assignment to the BARR model.

To further explore how the impact of BARR on this outcome varied across students, we also used the baseline NWEA MAP assessment data to divide the sample into four equal-sized groups. For this outcome and others, we found that this division yielded a similar pattern of subgroup findings, with impacts generally being larger for students scoring lower on the baseline NWEA assessment. For example, BARR reduced the incidence of course failure from 64.0 percent in the control group to 51.4 percent in the treatment group among those in the bottom quartile of the baseline NWEA achievement distribution. This compared to a reduction from 14.8 to 12.0 percent for those in the top quartile. See Appendix J for detailed findings from these analyses across a range of academic outcomes.

**Did BARR Impact Grade Point Average?**

In addition to examining BARR impact on course failure, we conducted an exploratory impact analysis on GPA. The GPA for BARR students was 2.58, compared to 2.48 for their counterparts in the control group (Exhibit 7). The difference of 0.10 grade points translates into an ES of 0.11. While closely tied to credit completion (and course failure), a student’s GPA is a broader measure of academic success that may capture variation hidden within a dichotomous pass/fail outcome. For example, BARR may improve the performance of students at the higher end of the achievement distribution, for whom course failure is a highly unlikely event.

The 0.11 ES shown in Exhibit 7 is smaller than the one we saw in Exhibit 6 for the course failure rate. This suggests that most of the effect on GPA occurred around the pass/fail cut point, which is consistent with the BARR model’s strong focus on preventing course failure, as observed during block/team meetings and especially risk-review meetings. This finding is also consistent with the pattern of subgroup findings we observed in Exhibit 6, which suggested that BARR is especially effective with students who are more likely to fail courses.

**Exhibit 7. Impacts on Students’ Grade Point Average in Core Courses (Full Sample and Subgroups)**

Outcome/Subgroup	N	BARR	Control	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>3,376</b>	<b>2.58</b>	<b>2.48</b>	<b>0.10‡</b>	0.000	<b>0.11</b>
Female	1,641	2.74	2.71	0.03	0.449	0.03
Male	1,735	2.43	2.26	0.17‡	0.000	0.17
Minority	2,497	2.48	2.35	0.14‡	0.000	0.13
White	879	2.88	2.86	0.03	0.566	0.03
Free or Reduced-Price Lunch	2,638	2.51	2.39	0.11‡	0.000	0.12
Not Free or Reduced-Price Lunch	722	2.90	2.80	0.10	0.053	0.11
English Learners	1,071	2.54	2.46	0.07	0.148	0.07

Outcome/Subgroup	N	BARR	Control	Difference	P-Value	Effect Size
Non-English Learners	2,305	2.61	2.48	0.12‡	0.000	0.12
Special Education	261	2.16	2.18	-0.02	0.853	-0.02
Non-Special Education	3,115	2.62	2.50	0.12‡	0.000	0.12

Source: AIR calculations from school-provided administrative data.

Notes. ‡ = statistically significant at the  $p < .001$  level.

Subgroup analyses of BARR’s impact on GPA show patterns similar to those we found for the course failure outcome in Exhibit 6. Male students, students of color, and students eligible for FRPL all had statistically significantly higher GPAs in the BARR group than in the control group. We also found positive impacts on GPA for students who were not English Language Learners and for students not classified as special education students.

**Did BARR Impact Reading and Mathematics Achievement?**

BARR did not have statistically significant impacts on the NWEA reading and mathematics test scores in the ten schools for which we had valid scores (Exhibit 8). Between 50 and 60 percent of students who completed baseline and follow-up NWEA assessments met their NWEA MAP growth projection, a rate that is consistent with NWEA expectations (Dahlin, 2013). Students assigned to BARR appeared somewhat more likely to meet their growth projection (ES = 0.10 for reading and ES = 0.06 for mathematics), but these impacts were not statistically significant.

**Exhibit 8. Impacts on Standardized Reading and Mathematics Scale Scores and Percentage Meeting Growth Projections**

Outcome	N	BARR	Control	Difference	P-Value	Effect Size
<b>Standardized NWEA Scores</b>						
Reading	2,595	219.86	219.73	0.13	0.737	0.01
Mathematics	2,715	227.94	227.78	0.16	0.688	0.01
<b>% Meeting Growth Projections</b>						
Reading	2,218	59.1	55.2	3.9	0.065	0.10
Mathematics	2,252	59.0	56.7	2.3	0.261	0.06

Source: AIR calculations from school-administered NWEA assessments.

Note. This table presents results for ten of the eleven study schools. One school was excluded from the NWEA outcome analyses due to difficulties encountered while administering the follow-up assessment.

BARR did not impact NWEA reading and mathematics scores for most demographic subgroups examined (Exhibits 9 and 11). One statistically significant difference was found for a subgroup,

indicating a small positive impact on mathematics scores for White students (ES = 0.09). There were more statistically significant subgroup impacts for the growth score outcomes (Exhibits 10 and 12). In reading, BARR students were more likely to meet their NWEA growth target than their control group counterparts if they were male (ES = 0.16), if they were not English learners (ES = 0.15), and if they were not special education students (ES = 0.12). In mathematics, BARR students were more likely to meet their NWEA growth target than their control group counterparts if they were male (ES = 0.17), if they were White (ES = 0.32), and if they were not English learners (ES = 0.17).

**Exhibit 9. Impacts on Standardized Reading Scale Scores, by Subgroup**

Outcome/Subgroup	N	BARR	Control	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>2,595</b>	<b>219.86</b>	<b>219.73</b>	<b>0.13</b>	<b>0.737</b>	<b>0.01</b>
Female	1,298	221.68	221.72	-0.05	0.925	0.00
Male	1,297	218.12	217.67	0.45	0.466	0.03
Minority	1,994	217.70	217.63	0.07	0.879	0.00
White	601	226.75	226.96	-0.21	0.780	0.01
Free or Reduced-Price Lunch	1,989	218.01	217.54	0.48	0.310	0.03
Not Free or Reduced-Price Lunch	595	226.21	227.00	-0.79	0.239	0.06
English Learners	844	214.44	215.25	-0.81	0.245	0.05
Non-English Learners	1,751	222.38	221.91	0.47	0.313	0.03
Special Education	197	205.83	206.18	-0.35	0.824	0.02
Non-Special Education	2,398	220.99	220.86	0.13	0.751	0.01

Source: AIR calculations from school-administered NWEA assessments.

Note. This table presents results for ten of the eleven study schools. One school was excluded from the NWEA outcome analyses due to difficulties encountered while administering the follow-up assessment.

**Exhibit 10. Impacts on Reading Growth Projections Met, by Subgroup**

Outcome/Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>2,218</b>	<b>59.1</b>	<b>55.2</b>	<b>3.9</b>	<b>0.065</b>	<b>0.10</b>
Female	1,115	58.3	56.9	1.4	0.631	0.04
Male	1,103	59.8	53.6	6.2*	0.04	0.16
Minority	1,642	58.2	55.0	3.2	0.193	0.08

Outcome/Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
White	576	60.3	56.5	3.8	0.354	0.10
Free or Reduced-Price Lunch	1,651	59.0	54.8	4.2	0.091	0.11
Not Free or Reduced-Price Lunch	556	59.7	56.0	3.7	0.366	0.10
English Learners	789	53.8	54.1	-0.3	0.947	0.01
Non-English Learners	1,429	61.6	55.9	5.7*	0.03	0.15
Special Education	169	47.6	53.5	-5.9	0.444	0.16
Non-Special Education	2,049	59.9	55.4	4.5*	0.041	0.12

Source: AIR calculations from school-administered NWEA assessments.

Notes. \* = statistically significant at the  $p < .05$  level. This table presents results for ten of the eleven study schools. One school was excluded from the NWEA outcome analyses due to difficulties encountered while administering the follow-up assessment.

**Exhibit 11. Impacts on Standardized Mathematics Scale Scores, by Subgroup**

Outcome/Subgroup	N	BARR	Control	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>2,715</b>	<b>227.94</b>	<b>227.78</b>	<b>0.16</b>	<b>0.688</b>	<b>0.01</b>
Female	1,349	227.33	228.04	-0.71	0.185	0.04
Male	1,366	228.49	227.55	0.94	0.106	0.05
Minority	2,113	224.52	224.77	-0.25	0.591	0.02
White	602	239.86	238.23	1.63*	0.018	0.09
Free or Reduced-Price Lunch	2,097	225.17	225.10	0.07	0.878	0.00
Not Free or Reduced-Price Lunch	606	237.63	237.11	0.52	0.475	0.03
English Learners	908	222.58	223.53	-0.95	0.158	0.05
Non-English Learners	1,807	230.52	229.94	0.58	0.230	0.03
Special Education	200	213.91	213.12	0.78	0.643	0.05
Non-Special Education	2,515	229.09	228.92	0.16	0.684	0.01

Source: AIR calculations from school-administered NWEA assessments.

Notes. \* = statistically significant at the  $p < .05$  level. This table presents results for ten of the eleven study schools. One school was excluded from the NWEA outcome analyses due to difficulties encountered while administering the follow-up assessment.

**Exhibit 12. Impacts on Mathematics Growth Projections Met, by Subgroup**

Outcome/Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>2,252</b>	<b>59.0</b>	<b>56.7</b>	<b>2.3</b>	<b>0.261</b>	<b>0.06</b>
Female	1,107	58.7	60.0	-1.3	0.658	0.04
Male	1,145	59.6	53.3	6.3*	0.03	0.17
Minority	1,672	55.7	56.8	-1.1	0.647	0.03
White	580	67.3	55.5	11.8†	0.004	0.32
Free or Reduced-Price Lunch	1,675	56.3	55.8	0.5	0.815	0.02
Not Free or Reduced-Price Lunch	565	66.7	59.7	7.0	0.082	0.19
English Learners	822	53.4	58.0	-4.6	0.184	0.13
Non-English Learners	1,430	61.9	55.8	6.1*	0.018	0.17
Special Education	163	52.9	54.8	-1.9	0.802	0.05
Non-Special Education	2,089	59.5	56.9	2.6	0.227	0.07

Source: AIR calculations from school-administered NWEA assessments.

Notes. \* = statistically significant at the  $p < .05$  level; † = statistically significant at the  $p < .01$  level. This table presents results for ten of the eleven study schools. One school was excluded from the NWEA outcome analyses due to difficulties encountered while administering the follow-up assessment.

When we introduced the NWEA assessments earlier, we noted that the overall quality of the NWEA follow-up assessments declined over time, with fewer students in Cohort 3 schools taking the assessment seriously and with greater differences between the BARR and control students in the rates of participation in the assessment. To better understand the consequences of these data problems, we conducted sensitivity analyses (presented in Appendix G). We also separately estimated BARR’s impacts on the NWEA reading and mathematics scores in each of the three cohorts. These impact estimates are shown in Exhibit 13.

**Exhibit 13. Impacts on Standardized Reading and Mathematics Scale Scores and Growth Projections, by Cohort**

Cohort	Outcome	N	BARR	Control	Difference	P-Value	Effect Size
<b>Standardized NWEA Scores</b>							
Cohort 1	Reading	862	221.77	220.82	0.96	0.160	0.07
	Mathematics	871	228.11	228.30	-0.19	0.798	0.01
Cohort 2	Reading	708	225.05	224.94	0.11	0.868	0.01
	Mathematics	713	236.42	234.18	2.24‡	0.001	0.12
Cohort 3	Reading	1,025	214.51	215.22	-0.71	0.293	0.04
	Mathematics	1,131	222.46	223.37	-0.91	0.141	0.05
<b>% Meeting Growth Projections</b>							
Cohort 1	Reading	595	61.5	54.5	7.0	0.075	0.19
	Mathematics	514	63.4	56.0	7.4	0.090	0.19
Cohort 2	Reading	689	61.9	55.0	6.9	0.073	0.18
	Mathematics	690	70.0	62.6	7.4*	0.047	0.20
Cohort 3	Reading	934	55.2	55.6	-0.4	0.902	0.01
	Mathematics	1,048	48.9	53.0	-4.1	0.179	0.11

Source: AIR calculations from school-administered NWEA assessments.

Notes. \* = statistically significant at the  $p < .05$  level; ‡ = statistically significant at the  $p < .001$  level. This table presents results for ten of the eleven study schools. One school was excluded from the NWEA outcome analyses due to difficulties encountered while administering the follow-up assessment.

The table shows that BARR had more positive impacts on NWEA mathematics test scores in the second cohort than in the first or third one. There were no statistically significant impacts on reading in any of the three cohorts.

**Impacts on Student Experiences**

We administered a survey to all ninth-grade students in BARR and control classrooms toward the end of the ninth-grade school year. These surveys assessed student experiences along six constructs: supportive relationships, expectations and rigor, engagement, sense of belonging, social and emotional learning, and grit. The findings are presented in the form of scale scores, which were derived mathematically as a summary of responses to multiple survey questions



within each construct. The data were mathematically centered so that scores above 50 are more positive and scores below 50 are more negative for each construct.

**Did BARR Change Students’ Experiences in School?**

BARR positively impacted student experiences, as reported on the student surveys in three areas (Exhibit 14). Compared to control group students, BARR students reported experiencing more supportive relationships in school (ES = 0.29), higher levels of teacher expectations and rigor in the classroom (ES = 0.25), and being more engaged in school (ES = 0.11). We did not find impacts on students’ sense of belonging in school, the extent of their social and emotional learning, or their grit.

**Exhibit 14. Impacts on Student Experiences Measured by Student Surveys**

Outcome	N	BARR	Control	Difference	P-Value	Effect Size
Supportive relationships	2,716	51.56	48.66	2.90‡	0.000	0.29
Expectations and rigor	2,745	51.34	48.86	2.48‡	0.000	0.25
Student engagement	2,721	50.57	49.51	1.06†	0.006	0.11
Sense of belonging	2,692	50.24	49.79	0.45	0.244	0.04
Social and emotional learning	2,700	50.08	49.93	0.16	0.683	0.02
Grit	2,685	49.99	50.01	-0.02	0.960	0.00

Source: AIR calculations from AIR-administered student surveys.

Note. † = statistically significant at the  $p < .01$  level; ‡ = statistically significant at the  $p < .001$  level.

Most student subgroups demonstrated statistically significant positive impacts on supportive relationships and expectations and rigor, with the exception of White students and students in special education (Exhibit 15). Similar results were found for student engagement. For that outcome, however, there were no statistically significant impacts for the subgroups of male and female students. The effects were especially pronounced for students of color and English learners.

**Exhibit 15. Impacts on Student Experiences From Student Surveys by Subgroup (Effect Sizes Only)**

Outcome/Subgroup	N <sup>a</sup>	Supportive Relationships	Expectations & Rigor	Student Engagement	Sense of Belonging	Social & Emotional Learning	Grit
<b>Full Sample</b>	<b>2,745</b>	<b>0.29‡</b>	<b>0.25‡</b>	<b>0.11†</b>	<b>0.04</b>	<b>0.02</b>	<b>0.00</b>
Female	1,356	0.31‡	0.16**	0.10	0.05	0.03	-0.05
Male	1,389	0.28‡	0.32‡	0.11	0.02	0.03	0.04
Minority	2,004	0.37‡	0.30‡	0.15†	0.03	0.03	0.03
White	741	0.11	0.11	0.02	0.07	0.03	0.07
Free or Reduced-Price Lunch	2,115	0.29‡	0.26‡	0.09*	0.04	0.02	0.01
Not Free or Reduced-Price Lunch	621	0.29‡	0.22**	0.19*	0.06	0.03	0.00
English Learners	872	0.36‡	0.39‡	0.15*	0.11	0.13	0.05
Non-English Learners	1,873	0.27‡	0.19‡	0.09*	0.03	0.03	0.03
Special Education	171	0.20	0.15	0.06	0.08	0.10	0.04
Non-Special Education	2,574	0.30‡	0.26‡	0.12†	0.04	0.01	0.01

Source: AIR calculations from AIR-administered student surveys.

Note. \* = statistically significant at the  $p < .05$  level; † = statistically significant at the  $p < .01$  level; ‡ = statistically significant at the  $p < .001$  level.

<sup>a</sup> Sample sizes varied slightly across survey outcomes because some students skipped some questions. This N represents the largest available sample size across the six outcomes shown in this table.

**Impacts on Student Behavior**

Exhibit 16 shows the impacts of BARR on three exploratory behavior outcomes: the percentage of students who were chronically absent (more than 10 percent of the time), the percentage of students who were ever suspended during the year, and the percentage of students who enrolled in 10th grade in the same district the following year.<sup>14</sup> (The inverse of this outcome is a conservative proxy for the dropout rate.) Note that the absenteeism outcome was available for only nine schools, and the suspension outcome was available for only seven schools. One school was not able to share suspension data, and three schools had so few students suspended that we were unable to use the outcome as part of our analysis. For most subgroups

<sup>14</sup> Some students included in the initial grade 9 rosters either never enrolled or transferred out of the study schools shortly after the beginning of the school year. To account for this, we excluded students who were enrolled less than 20 days in the study school from analyses of the grade 10 enrollment outcome. There was no significant difference in the proportion of students excluded after controlling for school fixed effects ( $p = 0.37$ ).cdf23ez

and outcomes, these analyses found no statistically significant impacts (see Appendix H for details).

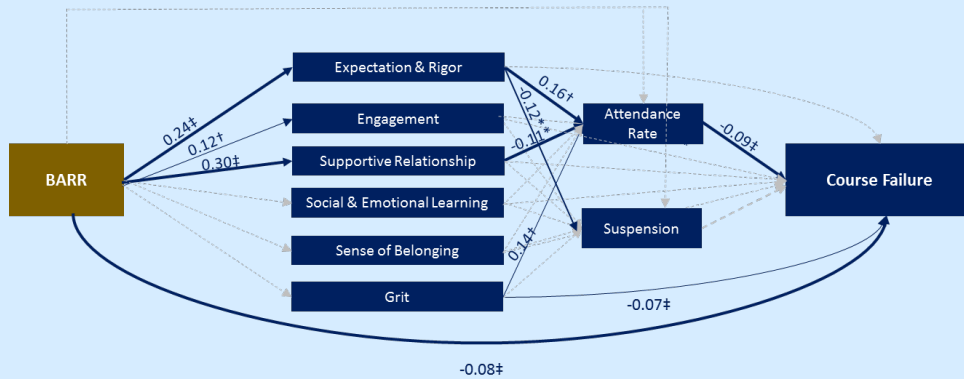
**Exhibit 16. Impacts on Students’ Chronic Absenteeism, Suspensions, and Persistence to Grade 10**

Outcome	N	BARR	Control	Difference	P-Value	Effect Size
Chronic absence (%)	3,275	23.1	21.9	1.2	0.314	0.04
Suspensions (%)	3,806	6.5	6.7	-0.2	0.774	0.02
Persistence to grade 10 (%)	2,863	88.8	87.4	1.4	0.222	0.08

Source: AIR calculations from school-provided administrative data.

**Evidence of Mediation**

A mediation analysis explores how impacts on proximal outcomes (e.g., student engagement) relate to impacts on distal outcomes (e.g., reductions in course failure). Understanding these relationships can help program developers and policymakers better understand and refine their interventions. We used a structural equation model (SEM), a type of correlational analysis, to explore the relationships between BARR and the outcome variables simultaneously. The figure below shows the results of this exploratory mediation analysis.



Model Fit Index: CFI = 0. 0.904; RMSEA = 0.063 (90% CI 0.058–0.068); SRMR = 0.035.

Notes. Blue arrows indicate significant direct relationships. (The bold blue arrows indicate the only mediation relationship that was found to be significant.) Gray arrows suggest insignificant relationships.

\* = statistically significant at the  $p < .05$  level; † = statistically significant at the  $p < .01$  level; ‡ = statistically significant at the  $p < .001$  level.

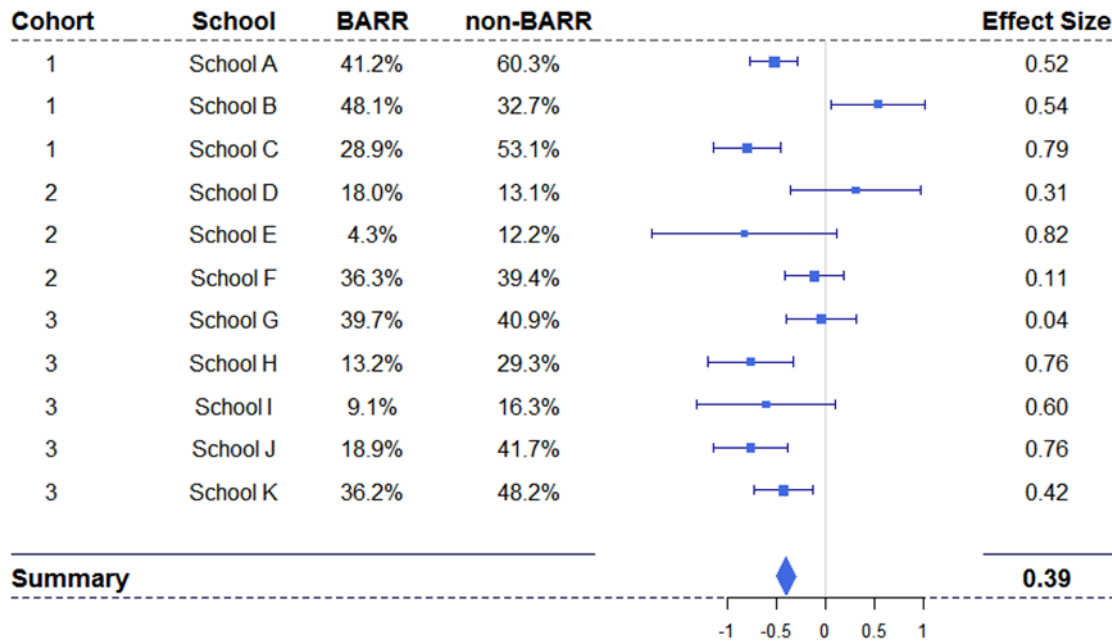
As shown in earlier exhibits, BARR had statistically significant favorable impacts on three of the six student survey constructs and on the rate of course failure experienced by students. The SEM analysis suggests that the positive impact on expectations and rigor was associated with a reduction in course failure via the intermediate outcome of student attendance. It also suggests that higher expectations and rigor were associated with a reduction in suspensions. As we saw in Exhibit 14, however, there was no statistically significant BARR impact on student attendance. The figure suggests that this may be the result of an apparent **negative** correlation between supportive relationships and attendance. That is, students with lower attendance reported having more supportive relationships with their teachers. As is sometimes the case in nonexperimental analyses such as these, it is possible that the relationships between those variables was reversed: Students with lower attendance may have sought out and may have needed more support from their teachers than students with higher attendance. Either way, the strong observed relationships between expectations, attendance, and course failure provide support for a continued emphasis on student attendance as a key mediator of academic success.

## How Did These Impacts Vary by School?

Because each participating high school in this study included both BARR and control group students, it is possible to estimate separate treatment effects for each school. The statistical precision of these estimated effects is limited and varies by school, with larger and more homogeneous schools having more precisely estimated impacts. In this section, we present the distribution of impacts across schools for the three confirmatory outcomes in the evaluation: the share of students failing one or more core courses, the NWEA reading assessment, and the NWEA mathematics assessment, including the growth projections (Exhibits 17 through 21). Each of these figures is a so-called “forest plot” in which each individual, school-level estimate is represented by a small box with two tails showing the 95-percent confidence interval surrounding the estimate. If neither of those tails crosses the vertical “zero” line, the impact estimate is statistically significant. The diamond at the bottom shows the weighted average of the individual school effects (i.e., the full sample estimate), and the width of the diamond shows its precision. Again, if no part of the diamond crosses the zero line, the overall impact estimate is statistically significant.

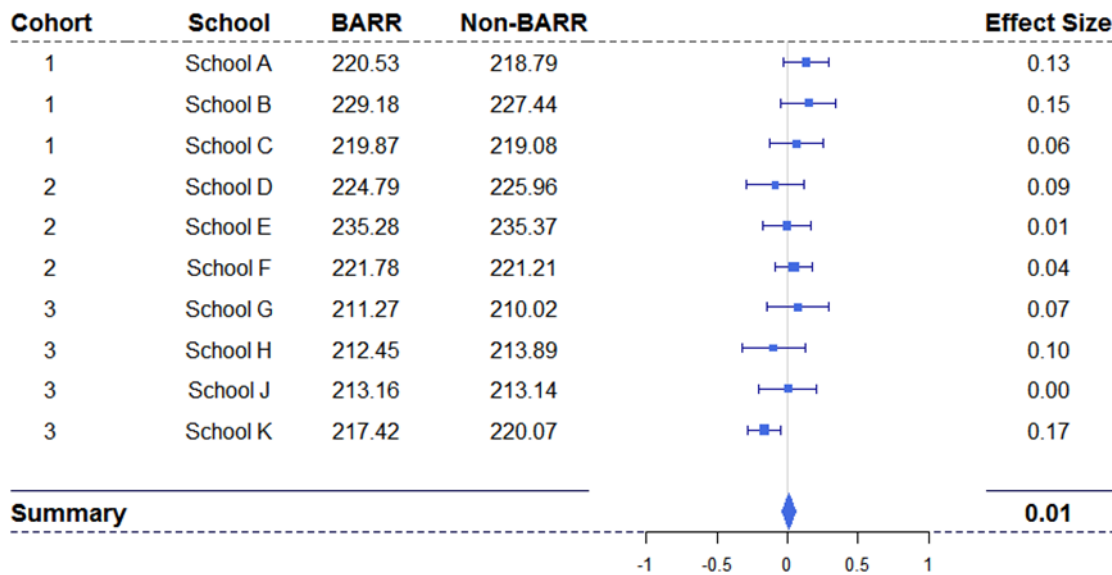
Most schools demonstrated favorable impacts on the course failure outcome (Exhibit 17). Only one school (School B in Exhibit 17) recorded an unfavorable impact estimate that was statistically significant. The largest favorable impacts were found in five Southern California schools. For NWEA reading and mathematics scores, most schools did not show significant differences between BARR and control group students (Exhibits 18 and 19). One school experienced negative impacts on both mathematics and reading, while two schools had positive impacts on mathematics. None of the school-level estimates for the percentage of students meeting their NWEA growth targets was statistically significant (Exhibits 20 and 21). Appendix I shows similar forest plots for all the other impact outcomes presented in this report.

**Exhibit 17. Estimated Effects on Students’ Failing One or More Core Courses, by School (11 Schools)**

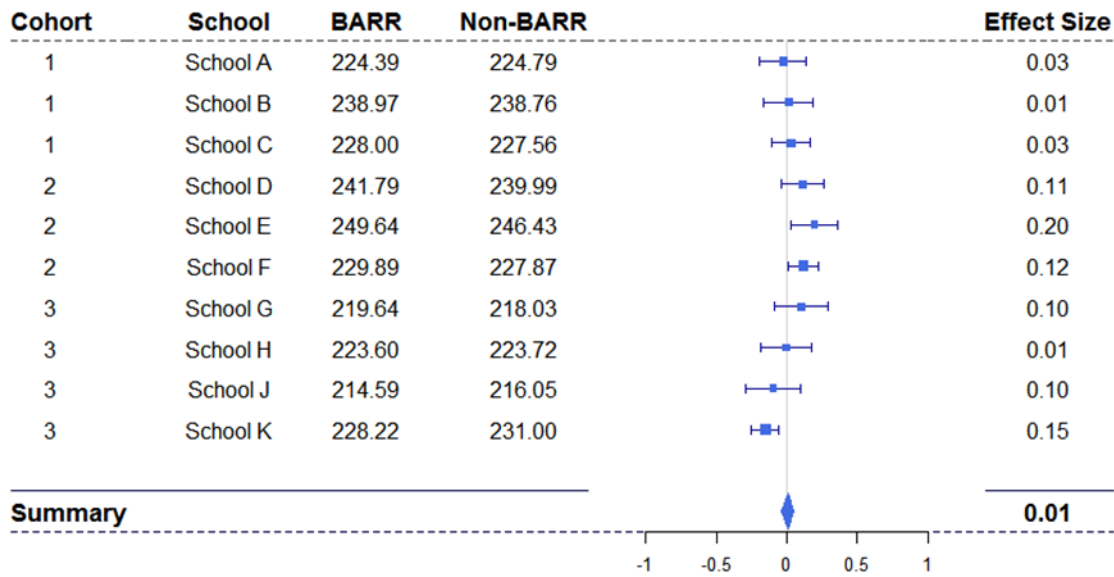


Note. Reduction in course failure is considered a favorable outcome for students.

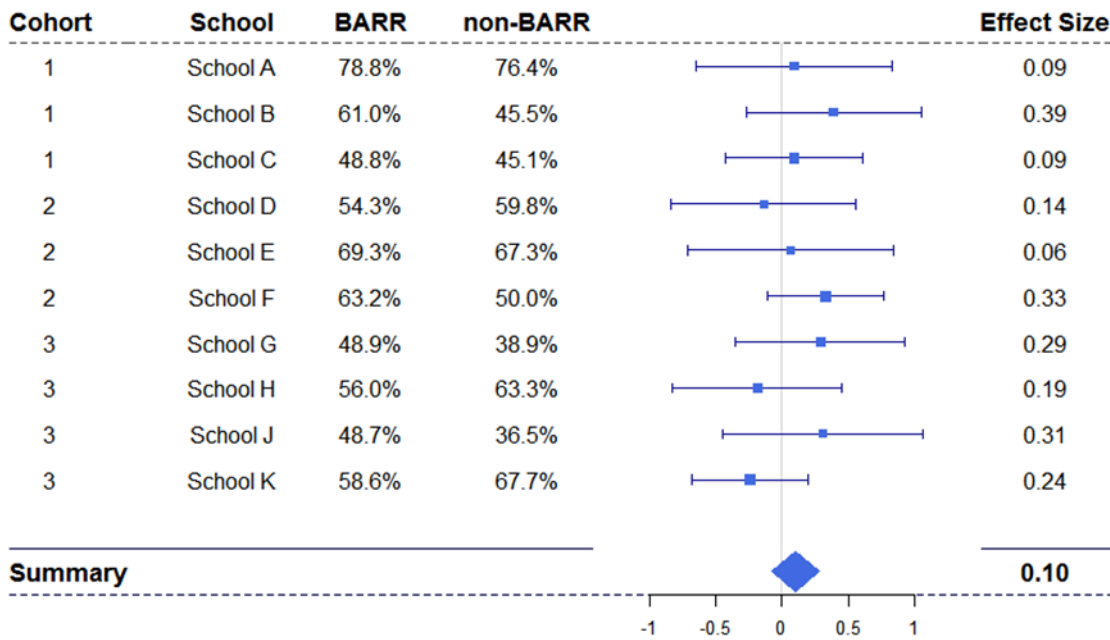
**Exhibit 18. Estimated Effects on Students’ Reading NWEA Scores, by School (10 Schools)**



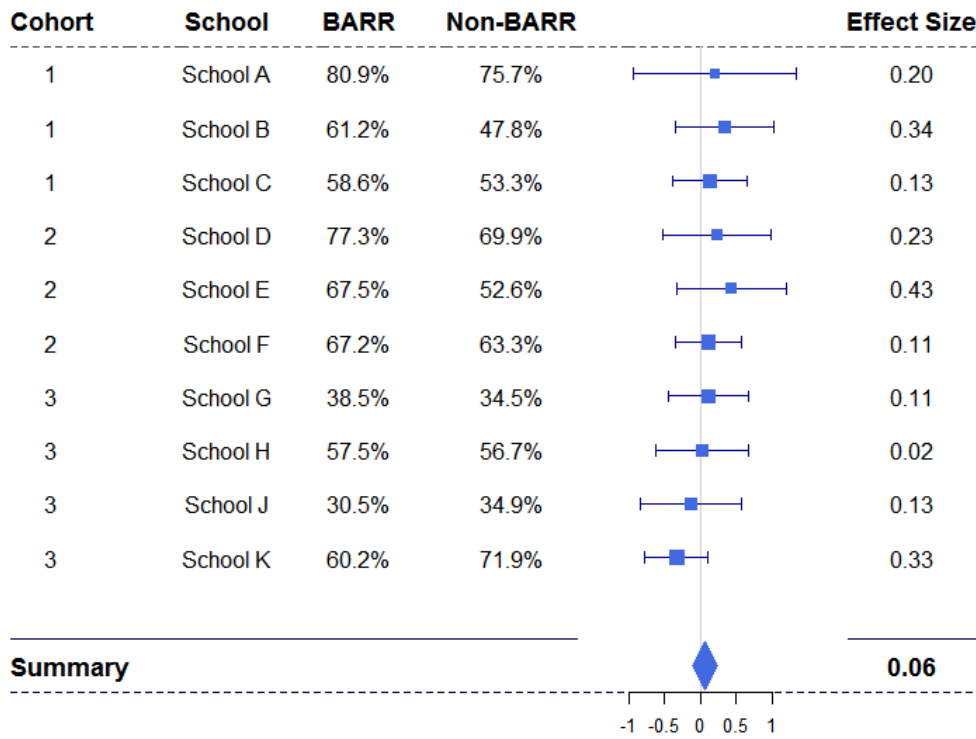
**Exhibit 19. Estimated Effects on Students’ NWEA Mathematics Scores, by School (10 Schools)**



**Exhibit 20. Estimated Effects on Students’ Reading NWEA Growth Projections Met, by School (10 Schools)**



**Exhibit 21. Estimated Effects on Students’ NWEA Mathematics Growth Projections Met, by School (10 Schools)**



**Experiences and Outcomes of the BARR Validation Schools Not Included in the Randomized Controlled Trial**

In addition to the randomized controlled trial (RCT), the findings of which are presented in this report, BARR was funded to disseminate the model to 35 mostly rural schools as part of the same i3 Validation Grant that funded this evaluation. The purpose of this dissemination work was to develop a broad adoption strategy and study implementation outcomes in these dissemination schools. In this box, we summarize some of the results of this dissemination effort. The findings presented here come from internal BARR research that did not involve AIR researchers.

**Design.** In the absence of an RCT, a school-level pre/post design with a dependent t-test was used to evaluate the effectiveness of the model with these 35 schools. Data on academic performance, attendance, and suspensions were obtained the year before BARR was implemented and after each year of implementation. Student and teacher surveys used for RCT validation schools were used in the dissemination schools. BARR evaluators also observed BARR implementation in these 35 schools and conducted interviews with BARR coordinators and principals.

**Sample.** The final dissemination school sample included 35 schools recruited in three cohorts: ten schools in 2015–16, fourteen in 2016–17, and eleven in 2017–18. These schools were located in Maine, Kentucky, North Carolina, Minnesota, and Wisconsin. Of these schools, thirteen were middle schools and 22 were high schools. A total of 33 schools were in rural districts. The dissemination schools included 16,070 students and 806 staff who were served during the grant period.

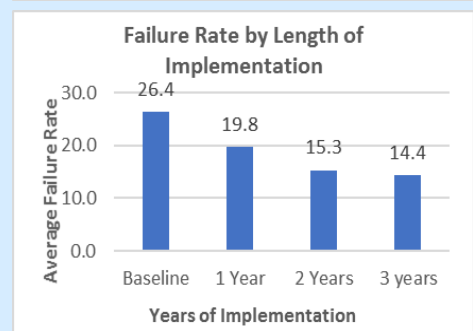
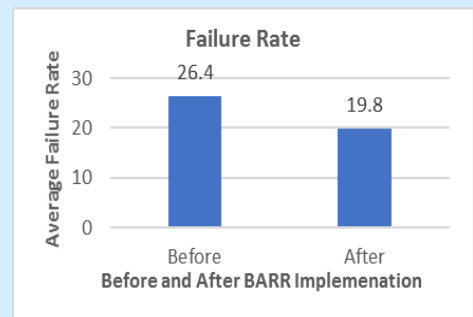
**Intervention.** Originally, the dissemination schools were supposed to receive a scaled-down version of the BARR model presented in this report, with fewer in-person visits, less coaching, and only a single year of BARR support. However, early in Year 1, BARR leadership determined that schools needed more than 1 year of implementation support and more frequent assistance for BARR to be sustainable in these schools. In addition to receiving two full years of support, dissemination schools received peer support at the BARR annual conferences and through the BARR Educator and BARR Administrator Networks. Also, while the dissemination schools received somewhat less hands-on support than the RCT schools, they did not have to deal with the challenges of within-school random assignment and partial implementation that faced RCT schools as a result of the evaluation design.

**Key Findings.** Descriptive data from the 35 dissemination schools indicated that all schools displayed improvement in at least one outcome measure (failure rate, absenteeism, or suspensions) in the first year of implementation. The longer a school implemented BARR, the lower the failure rate. Additional improvements were seen in attendance and behavior. Some schools expanded the BARR model into upper and lower grades, and all schools have continued to use at least some of the BARR strategies, if not the entire model, after the grant funding ended.

**Failure Rate.** Failure rate data were obtained for 11 of the 13 middle schools and 19 of the 22 high schools after the first year of implementation. Several middle schools implemented BARR in multiple grades. The average failure rate before BARR was 26.4%; after BARR, it was 19.8%. For schools that provided data, this 25% reduction in course failure was statistically significant ( $t(34) = 2.991, p < 0.005, ES = 0.5$ ).

The decrease in failure rate was 19% for middle schools and 31% for high schools. Overall, 86% of schools displayed a reduction in failure rate after 1 year. The failure rate continued to decrease the longer the school implemented BARR with an average failure rate of 14% by Year 3.

**Attendance.** Attendance was tracked using administrative data from the schools’ internal systems. Overall, 83% of schools reported improvements in either chronic absenteeism (missing 10% or more days), number of days absent, or average daily attendance. Among the seven schools reporting on chronic absenteeism, the average rate went from 22.9% before BARR to 19.1% after BARR, representing a 17% reduction in this outcome.





**Suspensions.** Sixteen schools reported the number of in-school and out-of-school suspensions. Before BARR, suspensions averaged 51.8 incidents; after BARR, suspensions averaged 44.1 incidents, representing an average reduction of 15% in the suspension rate. Overall, 75% of schools reported a decrease in suspensions.

**Implementation Fidelity.** Implementation fidelity was measured in the fall and spring in Year 1 using the same process described in this report for the RCT schools, and once in Year 2. A total of 94% of schools (33 out of 35) reached the threshold for adequate implementation fidelity by the end of Year 2. The major challenges for the two schools not meeting implementation fidelity were insufficient time for the BARR coordinator to receive coaching and inability to change the master schedule to allow for cohorts of students and teacher meeting time.

**Student and Teacher Surveys.** Student and teacher surveys were administered online during the spring, using the same instruments as for the RCT schools. For students, responses were highest for rigor and expectations, engagement, and supportive relationships; for teachers, responses were highest for collegiality, self-efficacy, and support from their administrators. (No statistical tests were done to compare these descriptive responses to one another).

**BARR Reflections for Recruitment, Implementation, Expansion, and Sustainability.** Working with the dissemination schools provided multiple opportunities to learn about recruitment, implementation, expansion, and sustainability of the BARR model. For recruitment to be successful, building and central office administrators and staff must all be involved. BARR schools benefited from early communication with coaches to develop the master schedule and prepare for implementation, as well as to make connections to other schools implementing BARR. Achieving implementation fidelity takes time and ongoing support from coaches, in addition to participation in BARR trainings in Years 2 and 3. Administrators reported that it is easier to start with one grade and expand rather than starting with multiple grades.

Supports to help schools implement BARR with fidelity included standardized training and supervision for coaches, tools to make coaching and evaluation more effective, annual trainings, an internet-based web tool for collecting data, videos that illustrate specific BARR components, and monthly professional learning communities for coordinators and coaches. To ensure that all students are succeeding, coaches provide support to schools on how to analyze school-specific data by relevant subgroups.

Many dissemination schools have expanded BARR to upper and lower grades, and administrators have spread BARR strategies, such as the strength-based approach, tracking failure rate, attendance, and parent/guardian contacts. As administrators saw the positive results achieved by BARR in high schools, they requested that BARR be implemented in middle schools, and thirteen middle schools were included in this study. Currently, there is a growing number of requests for an elementary school version of BARR. Several districts also see the advantage of having districtwide implementation of BARR to facilitate common beliefs, structures, and practices.

Schools have expressed a desire for ongoing support from the BARR Center after Year 3, citing the need to train new staff and address state mandates. On the basis of educator requests, BARR leadership has facilitated two networks to connect schools across the country. The BARR Educator Network shares best practices, solves problems, and provides support to BARR coordinators, while an Administrator Network tackles scheduling issues, data management and analysis, and provides support to principals and assistant principals. Both educators and administrators have presented their work at regional and national conferences. BARR has assisted schools in applying for federal, state, and private funding to provide additional support.

## Study Limitations

This study has several limitations. First, as mentioned, teachers were not randomly assigned to BARR. Assignment of teachers to implement the BARR program or “business as usual” was at the discretion of the schools participating in the study. Consequently, the effects of the program cannot be disentangled fully from those resulting from potential, underlying differences in teacher effectiveness. Although we did not find such differences in the measurable background characteristics that we collected from teachers, we cannot rule out that there were significant differences in unmeasured characteristics. When checking with principals about how they assigned teachers to the BARR program, we found that most assignments were driven by practical schedule considerations and that, typically, teachers neither were asked to volunteer nor were assigned to BARR as a professional development opportunity. However, we do not have complete or systematic data on this assignment process.<sup>15</sup>

A second limitation is that course completion (which is based on how teachers grade student work) is not a fully independent measure of student academic achievement. It is possible that teacher experiences in BARR and related improvements in teacher-student relationships may explain part of an effect on course completion in this evaluation. In our conversations with principals and teachers, however, we did not see any evidence of such indirect effects of BARR on teacher grading practices. If anything, both teachers and students in the BARR group reported that teachers had higher expectations of their students and held them to higher standards.

Third, as reported above, we encountered problems administering the NWEA MAP follow-up assessment, especially in the last cohort of study schools. The follow-up assessments were usually done very late in the school year, after all high-stakes testing had been completed. Even after excluding one school from the NWEA impact analysis and conducting sensitivity analyses to assess the effect of student attrition, there may be remaining concerns about how seriously students and schools took the NWEA assessments. This could affect the validity and reliability of the NWEA-based impact measures in this study, particularly given positive impacts on NWEA scores found in an earlier study of BARR (Corsello & Sharma, 2015).

Lastly, as discussed earlier, impacts may be attenuated due to the within-school random assignment and the fact that the experimental treatment contrast was limited to a single year of a 3-year intervention. Therefore, our impact estimates are a “partial dose” representation of the potential effect of the full BARR model.

---

<sup>15</sup> In one of the schools, the principal volunteered that he assigned his less effective core teachers to the BARR program in an attempt to improve their effectiveness. Because we did not collect similar information in all eleven schools, we did not act on this information or account for it in our impact analyses.

## Discussion and Conclusion

The findings presented in this report are an important contribution to the area of education research that focuses on high school effectiveness and the challenges that students face when transitioning into high school. This evaluation of the BARR model shows that high schools can significantly improve the experiences of their ninth-grade students and reduce their course failure rates by adopting a strength-based approach to reorganizing the ninth-grade class roster into blocks/teacher teams, encouraging teacher collaboration across academic departments, and creating a well-supported structure of tiered meetings to monitor the progress of all students and implement appropriate interventions for students who need them.

The evaluation found that the BARR model, originally developed for a single school, could be replicated successfully and implemented with fidelity in a variety of settings. This i3 validation study included large and small schools in a variety of states and in a range of communities and settings. Some of the schools in this study were rural, and others were in large cities. The rates of eligibility for free or reduced-price lunch ranged from 19.7 to 89.0 percent across the eleven schools, and there was similar variation in the percentage of students of color and the percentage of English learners.

Despite using a study design by which some ninth-grade students had BARR teachers for their core classes and others did not, this validation study was able to create and maintain a meaningful treatment contrast in all eleven schools. Across all schools and across a variety of implementation indicators, the experiences of the BARR students differed meaningfully from those of their counterparts in the business-as-usual control group. This included all eight components of the BARR model:

1. **Focus on the whole student.** BARR emphasizes the importance of acknowledging and mobilizing the assets that each student brings to school. At the end of ninth grade, BARR teachers had a significantly more positive view of their students' behavior than their counterparts in the control group. This finding was reflected in student surveys, in which BARR students reported that their teachers had higher expectations of them than of students in the control group.
2. **Professional development for teachers, counselors, and school leaders.** In all eleven schools, the national BARR Center was able to deliver summer professional development, monthly coaching, and ad hoc support to BARR teachers, BARR coordinators, and other school staff and leadership.

3. **I-Time curriculum.** This curriculum aims to increase the level of understanding that students and teachers have of one another’s circumstances, backgrounds, beliefs, assets, and challenges. It was implemented with fidelity in most schools. At the end of ninth grade, both students and teachers in the BARR group reported having stronger relationships with their peers and with one another than their counterparts in the control group.
4. **Create groups of students with common teachers.** In all eleven schools, the BARR model created distinct groups of students who shared the same set of core content teachers (ELA, mathematics, and science/social studies). These students took all their core classes together, which also may have contributed to students and teachers developing stronger relationships and students being less likely to “lose their way” during the ninth-grade school year.
5. **Regular block/team meetings to discuss and support student progress.** All but one school in the study implemented block/team meetings with fidelity. The sharing of student progress and of successful approaches to engaging or motivating students may have contributed to fewer students in the BARR group failing core courses during the school year.
6. **Risk-review meetings to address serious barriers to student success.** All but one school implemented risk-review meetings with fidelity. These meetings included teachers, BARR coordinators, school counselors, and other school staff as necessary.
7. **Engaging families.** Through the block/team meetings and risk-review meetings, teachers and other school staff were encouraged to engage parents in taking an active role in their children’s progress in school. Many interventions agreed upon during the block/team or risk-review meetings explicitly included parent engagement as a key component. Increased parent engagement also supported parents and school staff in better understanding and acknowledging the specific strengths and assets of each student (both in school and in the home environment).
8. **Engaging administrators.** In all but one school, the BARR team was able to engage administrators at the school and district levels as necessary for high-fidelity implementation of the BARR model. In addition to organizational and staff support, BARR teachers were more likely than their control counterparts to report having access to and actively using data to inform their instruction.

### ***Impacts of BARR***

The impact analysis found that BARR significantly improved student outcomes in several areas. First, students in the BARR group reported better relationships, greater teacher expectations, and greater engagement. Thus, their ninth-grade school experience was better than it would have been without BARR. Second, students in the BARR group were significantly less likely to fail any core courses in ninth grade and ended ninth grade with higher GPAs. These favorable impacts of BARR were stronger for students who traditionally do worse in ninth grade, including

students of color, male students, students who traditionally tested lower on standardized tests, and FRPL-eligible students.

We did not find statistically significant favorable or unfavorable impacts of BARR on NWEA test scores, a standardized measure of academic achievement administered in the spring of ninth grade. Thus, the impact of BARR on successful course completion, student engagement, and school climate may not yet have translated into an increase in academic skills as measured with this assessment. This contrasts with an earlier evaluation of BARR that did find such impacts. We also did not find statistically significant impacts on attendance, suspensions, and persistence into 10th grade.

The BARR model was not uniformly successful in all eleven validation study schools. We found significant cross-school variation in impacts. In general, impacts were strongest in large, diverse, urban and suburban schools. Students in the control group performed less well in these schools, thereby creating more room for improvement. In addition, maintaining the within-school treatment contrast may have been more difficult in small rural schools than it would have been in large urban schools. A larger study with more schools would be needed to systematically analyze variation in BARR's effectiveness across different school contexts.

In addition to the eleven-school impact study, the i3 validation grant enabled BARR to disseminate its model to 35 more schools. These schools received a less intensive version of the BARR model, but they did not have to maintain separate BARR and control groups. Internal evaluation research by the BARR developers found significant improvements over time across a similar range of outcomes, as we found in this RCT. These dissemination school findings therefore support the positive impact results we found in the RCT.

### ***Implications for the Field***

The BARR model fits squarely in a growing body of interventions targeting the climate in secondary schools, recognizing its impact on a range of student- and teacher-level outcomes, including teacher effectiveness, retention, student belonging, attendance, discipline, and academic progress. Unlike other reform models that can require dramatic changes to the school environment (e.g., changing school leaders and staff; breaking up large, comprehensive high schools into several smaller schools), BARR is unique in that it works within existing school structures. Aside from the once-a-week I-Time sessions, it does not introduce a new curriculum, and its staffing and training requirements are minimal compared to those of other interventions. Yet its impacts on the school environment, student experiences, and several academic outcomes are notable. Qualitative and quantitative data from the eleven-school impact study and the 35 dissemination schools show that the BARR model is popular in the schools in which it is used. The data suggest that bringing BARR into a school changes teachers' and administrators' ability and confidence to address their students'

needs from a whole-student perspective, including both the assets and the challenges that students bring with them. BARR teachers report better relationships with one another and with their students, and students similarly report stronger relationships with their teachers. Our data suggest that this interpersonal infrastructure may be as important for student success as the physical and environmental changes on which other successful interventions have relied. As the BARR model is more broadly scaled and adopted, it will be important to capture whether the success of this approach to high school reform can be sustained beyond the relatively small scale of this validation study, which still directly benefited from the time, commitment, and hands-on coaching of the original developer.

### **Conclusion**

The BARR model is a promising intervention to improve the experiences and outcomes of ninth-grade students. The intervention is modest in scope and effort and does not require major changes in curriculum, personnel, or school organization. BARR is more than an intervention designed to support at-risk students: It universally supports *all* students, with mechanisms in place to provide additional support to those most at risk. Although not directly assessed in this study, the cost of implementing BARR appears to be modest, both in terms of the necessary external support and in terms of time and resource requirements within participating schools. As a result, this model deserves consideration by schools and districts seeking to improve the ninth-grade transition in particular and school climate and teacher effectiveness in high school more broadly.

Our study leaves several open questions, most of which will be answered in the i3 scale-up study that is currently under way:

1. To what extent is the BARR model scalable beyond the point at which the original developer can engage with each participating school? What implementation supports are necessary in that case, and how intensive do these supports need to be?
2. Do BARR's academic impacts last beyond ninth grade, and do they translate into enduring impacts on student achievement, course-taking patterns, and graduation?
3. To what extent was the impact of BARR in this validation study affected by the fact that the program was not implemented schoolwide? Would BARR have had larger impacts if implementation had not been restricted to a randomly selected subset of ninth graders to accommodate this evaluation?

Results from the ongoing, 66-school i3 scale-up study will be available in 2021. In the meantime, we will continue to analyze data from this study to explore BARR's impacts on different subgroups of students, to examine variation in subgroup impacts across schools, and to describe the extent to which BARR's impact on academic outcomes was mediated by its impact on student experiences and student engagement.

## References

- Allensworth, E. M., & Easton, J. Q. (2005). *The on-track indicator as a predictor of high school graduation*. Chicago, IL: Consortium on Chicago School Research at the University of Chicago. Retrieved from <https://consortium.uchicago.edu/publications/track-indicator-predictor-high-school-graduation>
- Alvidrez, J., & Weinstein, R. S. (1993). The nature of “schooling” in school transitions: A critical re-examination. *Prevention in Human Services, 10*, 7–26.
- American Institutes for Research (AIR). (2012). *AIR conditions for learning surveys*. Washington, DC: Author. Retrieved from <https://supportiveschooldiscipline.org/resources/air-conditions-learning-surveys>
- American Institutes for Research (AIR). (2016). *The study of deeper learning: Opportunities and outcomes*. Washington, DC: Author. Retrieved from [https://www.air.org/sites/default/files/Deeper\\_Learning\\_Student\\_508-rev.pdf](https://www.air.org/sites/default/files/Deeper_Learning_Student_508-rev.pdf)
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561–573.
- Bandura, A. (n.d.). *Teacher self-efficacy scale*. Retrieved from <https://cpb-us-w2.wpmucdn.com/u.osu.edu/dist/2/5604/files/2014/09/Bandura-Instr-1sdm5sg.pdf>
- 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. Retrieved from <http://dx.doi.org/10.1111/j.1467-8624.2009.01265.x>
- Benson, P. L. (2007). Developmental assets: An overview of theory, research, and practice. In R. Silbereisen & R. Lerner (Eds.), *Approaches to Positive Youth Development* (pp. 33–58). Thousand Oaks, CA: SAGE.
- Bridgeland, J. M., Dilulio, J. J., & Morison, K. B. (2006). *The silent epidemic: Perspectives of high school dropouts*. Washington, DC: Civic Enterprises, LLC. Retrieved from <https://files.eric.ed.gov/fulltext/ED513444.pdf>
- Bruce, M., Bridgeland, J. M., Fox, J. H., & Balfanz, R. (2011). *On track for success: The use of early warning indicator and intervention systems to build a grad nation*. Washington, DC: Civic Enterprises, LLC. Retrieved from <https://files.eric.ed.gov/fulltext/ED526421.pdf>
- Cauley, K., & Jovanovich, D. (2006). Developing an effective transition program for students entering middle school or high school. *The Clearing House, 80*(1), 15–25.

- Christle, C. A., Jolivet, K., & Nelson, C. M. (2007). School characteristics related to high school dropout rates. *Remedial and Special Education, 28*(6), 325–339.
- Consortium on Chicago School Research (CCSR). (2007 version). *My Voice, My School: Student and teacher surveys*. Retrieved from (2015 version called 5 Essentials Survey) <https://consortium.uchicago.edu/sites/default/files/2018-10/Statewide%20E%20Report.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* (final report). Napa, CA: Corsello Consulting. Retrieved from <https://files.eric.ed.gov/fulltext/ED560804.pdf>
- Dahlin, M. (2013). *What percentage of students meet their growth targets?* Portland, OR: NWEA. Retrieved from <https://www.nwea.org/blog/2013/partner-questions-month-percentage-students-meet-growth-targets/>
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087.
- Dweck, C. S., Walton, G. M., & Cohen, G. L. (2014). *Academic tenacity: Mindsets and skills that promote long-term learning*. Seattle, WA: Bill & Melinda Gates Foundation.
- Easton, J. Q., Johnson, E., & Sartain, L. (2017). *The predictive power of ninth-grade GPA*. Chicago, IL: University of Chicago Consortium on School Research.
- Faria, A. M., Heppen, J., Li, Y., Stachel, S., Jones, W., Sawyer, K., ... Casserly, M. (2012). *Charting success: Data use and student achievement in urban schools*. Washington, DC: Council of the Great City Schools. Retrieved from [http://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Charting\\_Success.pdf](http://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Charting_Success.pdf)
- Goddard, Y. L., Goddard, R. D., & Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record, 109*(4), 877–896.
- Graham, P. (2007). Improving teacher effectiveness through structured collaboration: A case study of a professional learning community. *RMLE Online, 31*(1), 1–17.
- 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.



- Hartman, J., Wilkins, C., Gregory, L., Gould, L. F., & D'Souza, S. (2011). *Applying an on-track indicator for high school graduation: Adapting the Consortium on Chicago School Research indicator for five Texas districts* (Issues & Answers Report, REL 2011–No. 100). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from <https://www2.ed.gov/programs/slcp/2011progdirmtg/txontrackind.pdf>
- Henderson, A. T., & Mapp, K. L. (2002). *A new wave of evidence: The impact of school, family, and community connections on student achievement*. Austin, TX: Southwest Educational Development Laboratory.
- Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, 1(4), 85–108
- 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.
- Khmelkov, V. T., & Davidson, M. L. (2009). *Culture of excellence and ethics assessments: Description and theory*. Manlius, NY: Institute for Excellence and Ethics.
- Kraft, M. A., & Papay, J. P. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*, 36(4), 476–500.
- Lieberman, A. (Ed). (1990). *Schools as collaborative cultures: Creating the future now*. Bristol: Taylor & Francis.
- Levine, T. H. (2010). Tools for the study and design of collaborative teacher learning: The affordances of different conceptions of teacher community and activity theory. *Teacher Education Quarterly*, 37(1), 109–130.
- 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.

- Measures of Effective Teaching (MET) Project. (2012). *Asking students about teaching: Student perception surveys and their implementation*. Seattle, WA: Bill & Melinda Gates Foundation.
- Norbury, H., Wong, M., Wan, M., Reese, K., Dhillon, S., & Gerdeman, R. (2012). *Using the freshman on-track indicator to predict graduation in two urban districts in the Midwest Region* (Issues & Answers Report, REL 2012–No. 134). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Northwest Evaluation Association (NWEA). (n.d.). *Guidance for administering NWEA MAP/MPG assessments when results are used for high stakes purposes*. Portland, OR: Author. Retrieved from <https://www.nwea.org/content/uploads/2017/04/Guidance-for-Administering-MAP-and-MPG-Assessments-When-Results-are-Used.pdf>
- Panorama Education. (2015). *Validity brief: Panorama Student Survey*. Boston, MA: Author. Retrieved from <https://www.panoramaed.com/panorama-student-survey>
- Penuel, W. R., Sun, M., Frank, K. A., & Gallagher, H. A. (2012). Using social network analysis to study how collegial interactions can augment teacher learning from external professional development. *American Journal of Education*, 119, 103–36.
- Puma, M. J., Olsen, R. B., Bell, S. H., & Price, C. (2009). *What to do when data are missing in group randomized controlled trials* (NCEE 2009-0049). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from <http://ies.ed.gov/ncee/pdf/20090049.pdf>
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (Exp. ed.). Chicago, IL: University of Chicago Press.
- 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.
- 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.

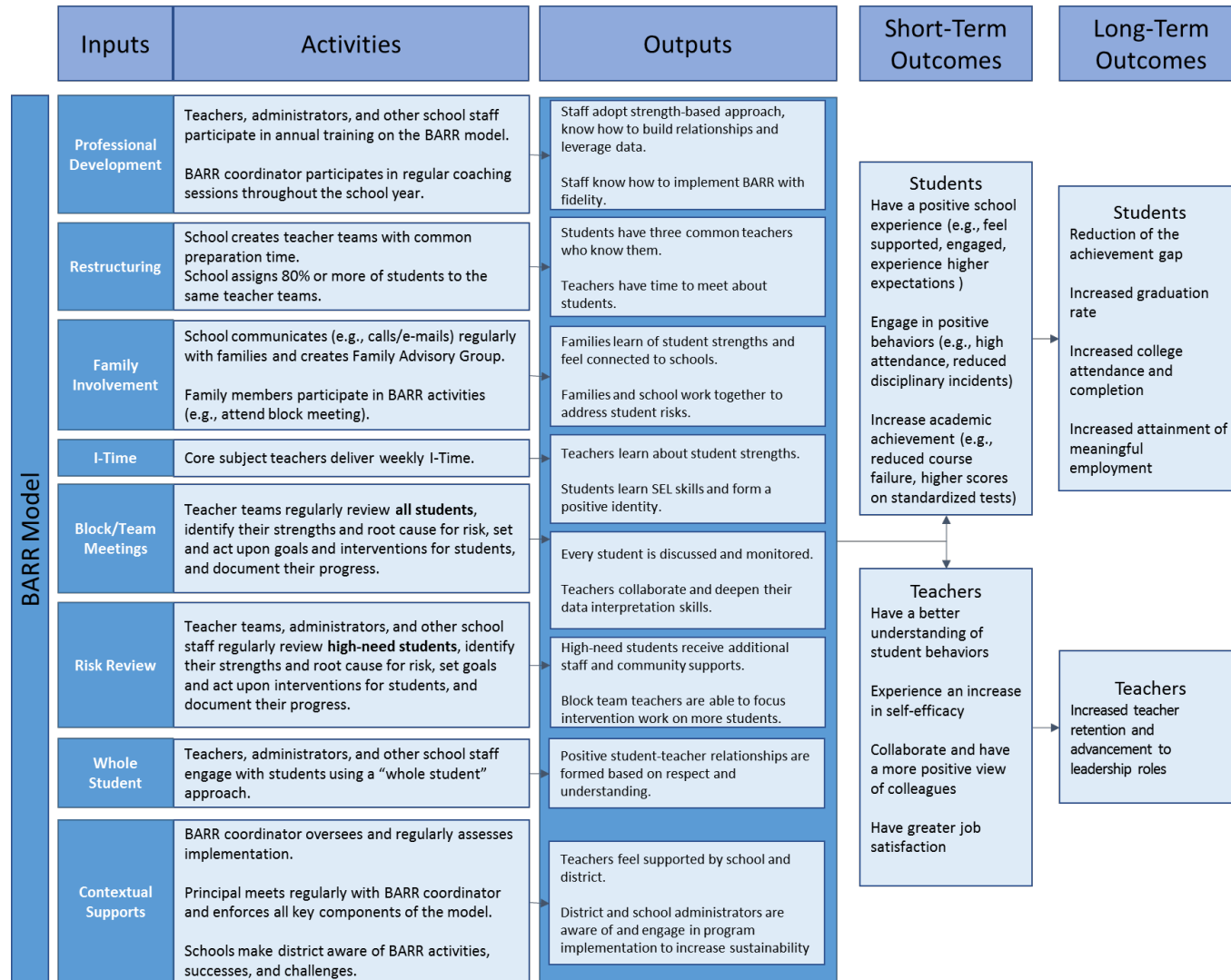
Stoker, G., Liu, F., & Arellano, B. (2017). *Understanding the role of noncognitive skills and school environments in students' transitions to high school* (REL 2018–282). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>

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

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

## Appendix A. BARR Logic Model

Appendix A details the BARR logic model.



## Appendix B. Data Collection Instruments

Appendix B includes the data collection tools for student academic outcomes, student experiences, and teacher experiences.

### B1. Student Survey Sources and Instrument

The student survey measured six student experience constructs and reliability of these survey scales met or exceeded acceptable internal consistency expectations (Cronbach’s alpha range from 0.77 to 0.89). Exhibit B1.1 documents the constructs, outcomes, data sources and reliability of survey scales.

#### Exhibit B1.1. Student Survey Constructs, Outcome Measures, Data Sources for Assessing the Impact of BARR on Student Experiences

Construct	Outcome Measures	Data Sources	Reliability (Cohort 1, 2, and 3)
Student Expectations and Rigor (Student Survey – Construct 1; SS1)	Students believe teachers have high expectations for their performance, provide clear guidelines, and encourage them to be successful.	Questions adapted from Panorama student survey, Tripod survey, CCSR my voice, my school surveys (CCSR, 2014; MET, 2012; Panorama Education, 2015). Scale consists of 11 items.	Cronbach’s alpha = 0.89
Student Engagement (SS2)	Students actively prepare for and participate in class activities, ask questions, and are interested in the lesson.	Questions adapted from Panorama study survey and Hewlett/AES surveys (AIR, 2016; Panorama Education, 2015). Scale consists of 7 items.	Cronbach’s alpha = 0.77
Supportive Relationships (SS3)	Students feel that their teachers are supportive of their interests and invested in their emotional well-being.	Questions adapted from Panorama student survey (Panorama Education, 2015). Scale consists of 7 items.	Cronbach’s alpha = 0.88

Construct	Outcome Measures	Data Sources	Reliability (Cohort 1, 2, and 3)
Social and Emotional Learning (SS4)	Students believe their classmates are capable of working through disagreements, managing their emotions, and acting responsibly.	Questions adapted from Conditions for Learning survey (AIR, 2012). Scale consists which consists of 11 items.	Cronbach’s alpha = 0.85
Sense of Belonging (SS5)	Students feel respected, accepted, and understood by their peers and that they belong with their classmates.	Questions adapted from Panorama student survey (Panorama Education, 2015). Scale consists of 4 items.	Cronbach’s alpha = 0.87
Grit (SS6)	Students are able to focus on and work toward meeting goals even when they encounter setbacks.	Questions adapted from Duckworth’s Grit Scale and Panorama student surveys (Duckworth, Peterson, Matthews, & Kelly, 2007; Panorama Education, 2015). Scale consists of 6 items.	Cronbach’s alpha = 0.88

The remaining content in this section details the questions aligned to each student experience construct. Information is formatted in a manner similar to what a student would experience during survey administration.

## Building Assets Reducing Risks Student Survey

*First we would like you to think about your English, social studies, math, and science classes.*

SS1. For how many of your **English, social studies, math, and science** classes is each statement **true**?

	None of my classes	One of my classes	Two of my classes	Three of my classes	Four of my classes
a. My teacher encourages me to do my best.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. My teacher <b>often</b> takes the time to make sure I understand the material	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. My teacher does not let people give up when the work gets hard.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. My teacher accepts nothing less than my full effort.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. My teacher pushes me to become a better thinker.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. My teacher makes us try to find the answers on our own before he or she answers our questions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. I learn a lot from feedback on my work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. The class-work helps me learn the course materials.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
i. The work we do in class is good preparation for tests.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
j. I know what my teacher wants me to learn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
k. It's clear what I need to do to get a good grade.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SS2. Still thinking about your **English, social studies, math, and science** classes, for how many of these classes is each statement true?

	None of my classes	One of my classes	Two of my classes	Three of my classes	Four of my classes
a. I <b>often</b> get so focused on class activities that I lost track of time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. I am interested in the class.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. When I am not in class, I <b>often</b> talk about ideas from class.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. I <b>often</b> practice in these classes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. I am <b>often</b> excited to go to my classes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. I <b>always</b> prepare for class.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. I ask questions when I do not understand the lesson.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Now we want you to think about your **English, social studies, math, and science** teachers.



SS3. Thinking about your **English, social studies, math, and science** teachers, for how many of those teachers is each statement true?

	None	One of my teachers	Two of my teachers	Three of my teachers
a. If my teacher asks me how I am doing, I <b>often</b> feel that they are really interested in my answer.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. My teacher is <b>interested</b> in my career after I finish school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. If I came back to visit class three years from now, my teacher would be excited to see me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. My teacher is <b>interested</b> in what I do outside of class.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. If I walked into class upset, my teacher would be <b>concerned</b> .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. If I had something on my mind, my teacher would <b>carefully</b> listen to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. I feel <b>connected</b> to my teacher.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Now we want you to think about the other students in your English, social studies, math, and science classes.

SS4. How much do you **agree or disagree** with each of the following statements about most students in your **English, social studies, math, and science** classes?

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. Stop and think before doing anything when they get angry.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Say mean things to other students when they think the other students deserve it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Give up when they can't solve a problem easily.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Think it's okay to fight if someone insults them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Try to work out their disagreements with other students by talking to them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Do all their homework.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Get into arguments when they disagree with people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. Try to do good job on school work even when it's not interesting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
i. Think it's ok to cheat if other students are cheating.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
j. Do their share of work when we have group projects.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
k. Do their best, even when their school-work is difficult.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SS5. Still thinking about the students in your **English, social studies, math, and science** classes this year, how much do you **agree or disagree** with each of the following statements?

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. I feel like I belong with the students in my classes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. I feel accepted by the students in my classes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. I feel like students in my classes understand me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Students in my classes show me respect.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**The last questions are about you.**

SS6. How often are the following statements **true** about you?

	Never/ Almost never	Sometimes	Usually	Always/ Almost always
a. I overcome setbacks to achieve important goals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. I do a careful and thorough job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. I finish what I begin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. I achieve goals even if they take a long time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. I am a hard worker.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. If I fail to reach an important goal, I will try again.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SS7. Is there anything else you'd like to add about your experience as a ninth grader this year? (Text box)

## B2. Teacher Survey Questions

The teacher survey measured eight teacher experience constructs related to teacher efficacy and effectiveness. Chronbach’s alpha for these survey scales ranged from 0.62 to 0.85. All but one scale, student accountability, met or exceeded acceptable internal consistency. Exhibit B2.1 documents the constructs, outcomes, data sources, and reliability of survey scales.

### Exhibit B2.1. Teacher Survey Constructs, Outcome Measures, and Data Sources for Assessing the Impact of BARR on Student Experiences

Construct	Outcome Measures	Data Sources	Reliability
School supports (Teacher Survey – Construct 1; TS1)	Teachers feel the school provides students and staff with necessary resources and support services.	Questions taken from AIR-developed survey to support Louisiana’s Safe and Supportive School Grant (AIR). Scale consists of 4 items.	Cronbach’s alpha = 0.79
Self-efficacy (TS2)	Teachers believe they have the ability and the power to affect learning, motivation, and behavior in classrooms.	Questions come from the Bandura’s Teacher Self-Efficacy Scale (Bandura, n.d.). Scale consists of 8 items.	Cronbach’s alpha = 0.83
Collaboration with and view of colleagues (TS3)	Teachers work together, trust each other, and have shared responsibilities and teaching approach.	A mix of newly created questions and questions from Bandura’s Teacher Self-Efficacy Scale (Bandura, n.d.). Scale consists of 8 items.	Cronbach’s alpha = 0.89
Data Use (TS4)	Teachers have access to data, discuss data, and use data to differentiate instruction for students.	Questions come from AIR’s Urban Data Study (Faria et al., 2012). Scale consists of 7 items (Cronbach’s alpha = 0.81 across all cohorts) collected on the spring teacher survey.	

Construct	Outcome Measures	Data Sources	Reliability
Perception of students' behavior, commitment, and attitudes (TS5)	What teachers think students would do to help their peers or make the right choices in a given situation.	Questions adapted from the Culture of Excellence and Ethics Assessment (Khmelkov & Davidson, 2009). Scale consists of 7 items.	Cronbach's alpha = 0.81
View of students' observed behavior, commitment, and attitudes (TS6)	What teachers observe students doing in classrooms related to goal setting, motivation, and completing work.	Questions adapted from the Culture of Excellence and Ethics Assessment (Khmelkov & Davidson, 2009). Scale consists of 8 items.	Cronbach's alpha = 0.82
Interaction with parents (TS7)	Teachers feel confident engaging with parents and helping them understand what students need to learn.	Questions adapted from the Culture of Excellence and Ethics Assessment (Khmelkov & Davidson, 2009). Scale consists of 6 items.	Cronbach's alpha = 0.85
Student accountability (TS8)	Teachers believe students are accountable for their own learning and grades and should meet all deadlines.	Newly created items for this study. Scale consists of 7 items.	Cronbach's alpha = 0.62

The remaining content in this section details the questions aligned to each teacher experience construct. Information is formatted in a manner similar to what a teacher would experience during survey administration.

## Building Assets Reducing Risks Teacher Survey

TS1. To what extent do you agree or disagree with the following statements about your school? This school...

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. Is a supportive and inviting place for students to learn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Provides effective counseling and support services for students.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Promotes trust and collegiality among staff.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Provides me with the materials, resources, and training I need to do my job effectively.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

TS2. How much can you do to...

	Nothing or very little	Some	A fair amount	A great deal
a. Overcome the influence of adverse community conditions on students' learning?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Promote learning when there is lack of support from the home?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Control disruptive behavior in the classroom?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Motivate students who show low interest in school work?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Get through to the most difficult students?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Get students to work together?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Keep students on task on difficult assignments?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. Get students to do their homework?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

TS3. To what extent do you agree or disagree with the following statements about the other Grade 9 teachers you work with?

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. We work together to coordinate our approach to addressing student needs.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. We meet regularly to discuss the performance of individual students.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. We meet regularly with counselors and school administrators to discuss the performance and needs of individual students.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. We trust each other.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. We really care about each other.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. We feel jointly responsible that all students learn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. We help maintain discipline in the whole school, not just our own classroom.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. I wish I had more time to meet with other teachers about my Grade 9 students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

TS4. To what extent do you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. I feel comfortable using data from my own Grade 9 classes to monitor and analyze my students' performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. 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.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. I have the tools I need to target specific interventions to my students when they need them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



	Strongly Disagree	Disagree	Agree	Strongly Agree
d. I do not have enough time to use the data I have effectively.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. I often talk with other teachers about performance and attendance data for my Grade 9 students.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. It is critical to have a complete picture of my students' performance to do my job properly.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. I have adequate support for effective use of available data on my students.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

TS5. To what extent do you agree or disagree about the following statements about most students in your Grade 9 classes? Most students in my Grade 9 classes...

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. Push themselves to meet high standards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Put off doing things they don't like to do.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Do just enough to get by on their schoolwork.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Set goals for doing better in school and keep track of whether they are improving.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Take pride in the quality of their work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Work hard to overcome their challenges.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Take initiative to get things done without being asked or reminded.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. Are open to suggestions for improvement.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

TS6. To what extent do you agree or disagree about the following statements about most students in your Grade 9 classes? Most students in my Grade 9 classes...

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. Are willing to help when they see someone having a problem.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Work well with students from different ethnic, religious, cultural, or political backgrounds.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Try to stop their friends from spreading rumors or gossip about others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Treat teachers and staff with respect, even when they disagree with them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Take an active role in helping solve school problems.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Encourage each other to follow the rules.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Help others on schoolwork, without letting them copy or cheat.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

TS7. To what extent do you agree or disagree with the following statements about the parents of your Grade 9 students?

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. I feel confident working with parents to help them support our students academically.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. I am able to help parents understand what social, emotional, and character skills our students need to learn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. I actively engage with parents when their child is having social, emotional, or behavioral challenges.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Strongly Disagree	Disagree	Agree	Strongly Agree
d. I often reach out to parents to let them know if their child has done something well or is making improvement.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. I have been able to show parents how to monitor their child's progress in school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. I am able to help parents find services in the community to support student needs.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

TS8. To what extent do you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. A student who does not complete an assignment on time should be allowed to turn in the assignment late.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. A student who fails an assignment or test should be given a chance to redo the assignment or retake the test.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. A student who is willing to complete alternative or extra credit assignments should be allowed those opportunities to improve their grade.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Giving students who are late with their assignments extra time is unfair to students who do their work on time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Allowing students to turn in assignments late, without penalty, leads to a decrease in their sense of personal responsibility.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. A student's grade should be based solely on test performance and the quality of assignments.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. A student's grade should be based in part on effort and class participation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. Which subject(s) are you currently teaching that involve Grade 9 students? (Select all)

- Math
- Science
- English language arts (reading, writing)
- Social studies/humanities
- Other (please specify)

10. What is your gender? (Select one)

- Male
- Female
- Prefer not to answer

11. What is your age range? (Select one)

- 18-21
- 22-25
- 26-30
- 31-40
- 41-50
- 51-60
- 61 and older

12. Including this school year, how long have you been teaching? (Select one)

- 1 year
- 2-3 years
- 4-5 years
- 6-10 years
- 11 years or more

13. Including this school year, how long have you been teaching at your current school? (Select one)

- 1 year
- 2-3 years
- 4-5 years
- 6-10 years
- 11 years or more
- Other (please specify)

14. What is the highest degree you have earned? (Select one)

- Associate's degree
- Bachelor's degree
- Master's degree Educational specialist diploma
- Ph.D., M.D., law degree, or other high-level professional degree
- I do not have a degree
- Teacher postsecondary course-taking

15. Would you recommend teaching at your school to a friend or graduate from your college? (Select one)

- Yes
- No

16. Is there anything else you'd like to add about your experience teaching ninth grade this year? (Text box)

## B3. Teacher Interview Questions

### B3.1. BARR Teacher Interview (Cohort 1, Treatment)

Hello, my name is \_\_\_\_\_ and I work for the American Institutes for Research. 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 evaluation and your rights as a participant. Do you have any questions before we get started?

I would like to be able to tape-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 what your personal experience has been like 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?*

##### 1) *Teacher outcomes*

- *Increased teacher effectiveness*
- *Being prepared to work with the ninth graders*
- *Increased teacher collaboration*
- *Improved staff relationships*
- *Feeling supported*

##### 2) *Building positive intentional relationships*

- *Staff-to-staff*
- *Staff-to-students*
- *Students-to-students*

##### 3) *Using real-time student-level performance data to guide instructional action*

We will now ask some more detailed questions related to each of the eight BARR strategies. You will then be asked to reflect on your experience and provide a 1 to 5 rating on the impact you think each strategy had on you and your effectiveness as a teacher.

**Impact Rating Scale to Prompt Respondents**

Complete impact	5	Completely transformative of your effectiveness as a teacher from last year.
Major impact	4	Major impact on your effectiveness as a teacher. (Most of the time)
Moderate impact	3	Moderate impact on your effectiveness as a teacher. (Sometimes)
Minimal impact	2	Minimal impact on your effectiveness as a teacher. (Rarely)
No impact	1	No impact at all on your effectiveness as a teacher. Unchanged from last year.

**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?*

- *The foundational training on the BARR theory and practices*
- *Periodic training on identified student needs*
- *Monthly meetings with BARR teachers*

5. Were there any particular aspects of this professional development that you think helped you feel successful as a teacher this year? Please provide examples.
6. In what ways, if any, did this professional development help you:
  - a. *Feel prepared to work with ninth-grade students?*
  - b. *Understand your students’ needs and assets?*
7. How much do you feel this professional development has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.
  - a. *What about the BARR-specific professional development? Is it the same or a different rating? [1-5 Rating] Please explain why.*

**C. Restructuring Ninth Grade**

8. Please describe some of the ways in which ninth-grade teaching and learning was restructured this year by the implementation of the BARR model at your school, if at all.

- *Smaller class size*
- *Core teachers work as a team*
- *Students share teachers across core classes*
- *Team has a common planning period*
- *Assigned a counselor for the ninth grade*

9. Were there any particular aspects of the BARR ninth-grade restructure that you think helped you feel successful as a teacher this year? Please provide examples.
10. What impact, if any, has the BARR ninth-grade restructure had on helping you:
  - a. *Feel prepared to work with ninth-grade students?*
  - b. *Develop personal and connected relationships with students?*
11. How much do you feel the BARR ninth-grade restructure has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

#### D. Block/Team Meetings

12. Please describe some of the practices you use (alone or with other teachers) to monitor and support students.
  - a. *In what ways, if any, have you used real-time student-level performance data to guide instructional decisions?*

- *Reach consensus for the agreed-upon goals and action plans*
- *Identify root causes and strengths/assets for high-need students*
- *Set individualized goals for each student*
- *Implement an action plan with appropriate interventions*
- *Refer high-need students to Risk Review*

13. Specific to BARR, were there any particular aspects of the block/team meetings that you think helped you feel successful as a teacher this year? Please provide examples.
  - a. *[IF APPLICABLE] In what ways, if any, is this different from how you felt about the team structure that was at your school previously?*

14. In what ways, if any, do you feel the BARR block/team meetings have helped you:
  - a. *Collaborate with other teachers?*

- *Feel like your decisions are supported by your block/team*
- *Assisted other teachers in making instructional decisions and addressing student learning barriers*
- *Have more people to talk to now than you did before*
- *Go to different people now than you did before*

- b. *Build positive peer-to-peer relationships?*

- *Feel respect and trust in your block/team*
- *Feel comfortable discussing feelings and frustrations with your block/team*

15. How much do you feel the BARR block/team meetings have impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.



### E. Whole Student Approach

[If prompt needed for Question 17: By “Whole Student approach” we are talking about working with the student as a whole person (i.e., academically, emotionally, physically, and socially), and looking at a student’s overall assets and strengths as well as risk factors.]

16. In what ways do you think a student's academic and nonacademic needs and strengths are expressed in the classroom?
  - a. *Please describe some of the practices you have used this year to address students’ needs and leverage their strengths, if any.*
  - b. *Can you give me an example of how you have helped a student with a need or a strength recently, say in the last month or so?*
17. Specific to BARR, were there any particular practices related to the Whole Student approach that you think helped you feel successful as a teacher this year? Please provide examples.
18. How much do you feel the way you addressed student needs and strengths has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

### F. I-Time

19. Were there any particular aspects of I-Time that you think were helpful to you as a teacher this year? Please provide examples.

- *Classroom management*
- *Positive youth development and assets*
- *Knowledge of students’ assets and challenges*
- *Encouraging student self-exploration*
- *Teacher seen as approachable*
- *Content developmentally and culturally competent*
- *Classroom spirit*

20. In what ways, if any, do you think the I-Time sessions have impacted your approach to teaching?
  - a. *Related to developing staff-to-student relationships?*
  - b. *Related to developing student-to-student relationships?*

- *Feel better prepared to work with ninth-grade students*
- *Improve your classroom management*
- *Understand positive youth development*
- *Understand students’ assets and challenges*
- *Have time for individualized instruction and support for students*
- *Develop more personal and connected relationships with students*
- *Set greater performance demands on students*

21. How much do you feel having I-Time has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

## G. Risk Review

22. Related to referring students to Risk Review, how do you think this compares to the previous RTI or PLC process?

a. *What is the level of feedback you receive for your referred students?*

23. In what ways, if any, do you feel the Risk Review process has impacted your ability to support high-needs students?

- *Understand the full range of interventions and resources available to high-needs students*
- *Feel more empowered to support high-need students*
- *Feel relief not identifying and supporting high-need students on your own*

24. How much do you feel the BARR Risk Review process has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

## H. Parent Involvement

25. Please describe some of the practices or activities you have used this year to try and foster positive parent-teacher relationships.

- *Parent orientation at beginning of school year*
- *Communication to parents (e.g., phone calls, e-mails)*
- *Parents participation in block/team meetings and risk review meetings*

26. What impact, if any, have you seen on parent involvement this school year compared to last year?

- *Frequency of contact with parents increased*
- *Parents able to collaborate with teachers more*
- *More opportunities and resources to support parents (i.e., role modeling and coaching)*
- *Feedback from parents about aspects of teaching and learning*

27. How much do you feel your practices or activities related to parent involvement have impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

## I. Contextual Support

28. What kinds of support, if any, were provided by the school leadership this year that helped your planning and instruction? Please provide examples.

a. *What kinds of support, if any, were provided by school leadership for the BARR implementation?*

- *Supports teacher decisions*
- *Expresses an awareness of issues with 9th-grade students*
- *Attends BARR trainings and meetings and is knowledgeable of the program*
- *Actively helps implement key BARR components in the school*
- *Advocates for sustaining and expanding BARR across grades and/or schools*

29. In regards to the role of the BARR coordinator, what kinds of support, if any, were provided that helped you this year? Please provide examples.
30. How much do you feel the support from school leadership this year has impacted your effectiveness as a teacher? [1-5 Rating] Please explain why.

We now have a few wrap up questions about your overall experience with BARR this year.

**J. Overall Experience with BARR**

31. Whether in a formal BARR structure or not, do you think you will continue with some of these practices next year? Why or why not?
  - a. *If so, what supports would help you to continue with these practices?*
32. Were there any aspects of the BARR strategies that you found confusing or difficult to integrate into your teaching?
  - a. *Were there any BARR topics or practices that you would have liked more related PD or support?*
33. Were there any major challenges or barriers to effective teaching at your school that you ran into this year?
  - a. *Were there any major challenges or barriers specific to implementing BARR at your school this year?*
34. To what extent, if any, do you think information and practices related to BARR were shared with other (non-BARR) teachers outside of your own BARR team?
  - a. *If so, what information do you think was shared with other teachers about BARR?*
35. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

**B3.2. NonBARR Teacher Interview (Cohort 1, Control)**

Hello, my name is \_\_\_\_\_ and I work for the American Institutes for Research. I am here today as part of an evaluation of the implementation of a pilot program at your school this year.

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

This discussion should take about 30-45 minutes. Please take a moment to review the attached interview consent form for more detailed information about this evaluation and your rights as a participant. Do you have any questions before we get started?

I would like to be able to tape-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 what your personal experience has been like teaching ninth graders this year?

For this interview, we will ask some questions related to a selection of teaching approaches and strategies. You will then be asked to reflect on your experience and provide a 1 to 5 rating on how much you think your effectiveness as a teacher has changed this year.

Complete impact	5	Completely transformative of your effectiveness as a teacher from last year.
Major impact	4	Major impact on your effectiveness as a teacher. (Most of the time)
Moderate impact	3	Moderate impact on your effectiveness as a teacher. (Sometimes)
Minimal impact	2	Minimal impact on your effectiveness as a teacher. (Rarely)
No impact	1	No impact at all on your effectiveness as a teacher. Unchanged from last year.

**B. Professional Development**

4. What professional development opportunities have you participated in this past year?
5. Were there any particular aspects of this professional development that you think helped you feel successful as a teacher this year? Please provide examples.
6. In what ways, if any, did this professional development help you:
  - a. Feel prepared to work with ninth-grade students?
  - b. Understand your students’ needs and assets?

7. How much do you feel this professional development has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

### C. Ninth-grade Structure

8. Please describe the current ninth-grade teaching and learning structure at your school. (i.e., class schedules, planning times)
9. What impact, if any, do you think the current ninth-grade structure has had on helping you:
  - a. *Feel prepared to work with ninth-grade students?*
  - b. *Develop personal and connected relationships with students?*
10. How much do you feel the current ninth-grade structure has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

### D. Student Progress Meetings

11. Please describe some of the practices you use (alone or with other teachers) to monitor and support students.
  - a. *In what ways, if any, have you used real-time student-level performance data to guide instructional decisions?*
12. Are you currently part of a team structure at your school where you share the same students across classes and are able to have regular team meetings?
  - a. *[If yes:] How often do you meet with your team?*
  - b. *What is the process for talking about students on your team?*
  - c. *Are you able to discuss students' needs and strengths that may be expressed across different classrooms?*
  - d. *Are you able to discuss possible instructional decisions with your colleagues?*
13. In what ways, if any, do you feel meeting with your team/other colleagues has helped you:
  - a. *Collaborate with other teachers?*
  - b. *Build positive peer-to-peer relationships?*
14. How much do you feel meeting with your team/other colleagues has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

### E. Whole Student Approach

15. In what ways do you think a student's academic and nonacademic needs and strengths are expressed in the classroom?
  - a. *Please describe some of the practices you have used this year to address students' needs and leverage their strengths, if any.*
  - b. *Can you give me an example of how you have helped a student with a need or a strength recently, say in the last month or so?*
16. How much do you feel the way you addressed student needs and strengths has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

## F. Social Emotional Learning/Youth Development

17. Are there any particular social and emotional learning or youth development practices/activities/curriculum that you used in your classroom this year? If so, Please provide examples.
18. In what ways, if any, do you think these social and emotional learning or youth development practices/activities/curriculum have impacted your approach to teaching?
  - a. *Related to developing staff-to-student relationships?*
  - b. *Related to developing student-to-student relationships?*
19. How much do you feel these social and emotional learning or youth development practices/activities/curriculum have impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

## G. RTI or PLC Process

20. What is the current RTI or PLC process for referring high-needs students for supports and services?
  - a. *What is the level of feedback you receive for your referred students?*
21. In what ways, if any, do you feel the current RTI or PLC process has impacted your ability to support high-needs students?
22. How much do you feel the current RTI or PLC process has impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

## 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.
24. What impact, if any, have you seen on parent involvement this school year compared to last year?
25. How much do you feel your practices or activities related to parent involvement have impacted your effectiveness as a teacher this year? [1-5 Rating] Please explain why.

## I. Contextual Support

26. What kinds of support, if any, were provided by the school leadership this year that helped your planning and instruction? Please provide examples.
27. How much do you feel the support from school leadership this year has impacted your effectiveness as a teacher? [1-5 Rating] Please explain why.

We now have a few wrap up questions about your overall teaching experience this year.

## J. Overall Experience

28. Were there any major challenges or barriers to effective teaching at your school that you ran into this year?
29. To what extent, if any, have you discussed information and practices with teachers involved in the pilot program?
  - a. *If so, what information have you heard about the pilot program?*

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

Thank you for your time!

### **B3.3. BARR Teacher Interview (Cohorts 2 and 3, Treatment)**

Hello, my name is \_\_\_\_\_ and I work for the American Institutes for Research. 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 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.
  - a. *Did anything specifically help prepare you to work with ninth-grade students?*
7. Were there any aspects of this professional development that you found confusing or difficult to integrate into your teaching?
  - b. *Were there any topics or practices for which you would have liked more PD?*

### 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.*
10. Were there any particular aspects of this ninth-grade structure that you think helped you as a teacher this year? Please provide examples.

### D. Team Meetings

11. 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?*
12. 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?*
13. 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.*
14. 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)*
15. *Are there any aspects of team meetings that you found challenging or difficult this year? Please provide examples.*

### E. Whole Student Approach

16. 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)
17. 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.
18. 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**

- 19. Did you facilitate any social and emotional learning or youth development activities/lessons in your classroom this year? (Including I-Times, but also probe for any other 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.
- 20. To what extent do you feel using social and 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.
- 21. In what ways, if any, do you think incorporating social and 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**

- 22. Please describe your experience with the current process for referring high-needs students for supports and services (e.g., Risk Review, Student Assistance Team (SAT)).
  - a. *If you have referred students, what is the level of feedback or follow-up you have received from this process?*
- 23. 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**

- 24. 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?*
- 25. To what extent do you feel these parent involvement practices or activities 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.
- 26. To what extent have you seen actual parent involvement change this school year compared to last year? (*Better, the same, worse?*)

#### **I. Contextual Support**

- 27. 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 the BARR implementation?*

28. 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.
29. 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?*
30. 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**

31. Compared to ninth-grade students you may have had in the past, what do you think the school experience has been like for this 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?*)
32. 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)*
33. To what extent, if any, do you think information and practices related to BARR were shared with other (non-BARR) teachers outside of the BARR team(s)?
  - a. *If so, what information do you think was shared with other teachers about BARR?*
34. Thinking about your overall experience this year with BARR, is there something that stands out above everything else? Please explain.
35. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

### **B3.4. NonBARR Teacher Interview (Cohorts 2 and 3, Control)**

Hello, my name is \_\_\_\_\_ and I work for the American Institutes for Research. I am here today as part of our work evaluating the implementation of initiative pilot program at your school 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.
  - a. *Did anything specifically help prepare you to work with ninth-grade students?*
7. Were there any aspects of this professional development that you found confusing or difficult to integrate into your teaching?
  - a. *Were there any topics or practices for which you would have liked more PD?*

#### **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*)
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.
- 10. Were there any particular aspects of this ninth-grade structure that you think helped you as a teacher this year? Please provide examples.

#### **D. Team Meetings**

- 11. 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?*
- 12. 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?*
- 13. 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.
- 14. 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)*
- 15. *Are there any aspects of team meetings that you found challenging or difficult this year? Please provide examples.*

#### **E. Whole Student Approach**

- 16. 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)
- 17. 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.
- 18. 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**

- 19. Did you facilitate any social and 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.

20. To what extent do you feel using social and 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.
21. In what ways, if any, do you think incorporating social and 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

22. **Please describe your experience with the current process for referring high-needs students for supports and services (e.g., Student Assistance Team (SAT)).**
  - a. *If you have referred students, what is the level of feedback or follow-up you have received from this process?*
23. **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

24. 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?*
25. To what extent do you feel these parent involvement practices or activities 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.
26. To what extent have you seen actual parent involvement change this school year compared to last year? (*Better, the same, worse?*)

### I. Contextual Support

27. What types of support, if any, did the school administration provide to you that helped you this year? Please provide examples.
28. 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.
29. 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.
30. 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

31. Compared to ninth-grade students you may have had in the past, what do you think the school experience has been like for this 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?*)
32. Were there any major external or internal challenges that impacted your school this year?
33. To what extent, if any, have you discussed information and practices with teachers involved in the pilot program? (*Note: If they have heard of BARR, it is ok to ask specifically about BARR*)
  - a. *If so, what information have you heard about the pilot program/BARR?*
34. If your school decides to roll out BARR to the rest of the school next year, would you be interested in participating? Why or why not?
35. Is there anything you would like to add that I didn't ask about?

Thank you for your time!

#### **B4. Description of NWEA Tests**

Northwest Evaluation Association’s (NWEA) Measures of Academic Progress (MAP) Reading and Mathematics tests are Common Core aligned, computer adaptive achievement assessments designed to measure student achievement and growth over time. The tests can be administered 2 to 3 times during the school year—the fall, spring, and an optional test date in the winter. NWEA uses a RIT (Rasch unit) scale to generate a standardized interval score. A RIT score indicates the difficulty level at which the student is answering about 50% of the questions correctly. Although the test is not typically timed, developers state that students spend approximately 60 minutes per subject area. For the MAP test, anything shorter than 15–20 minutes in duration can be associated with inaccurate estimates (NWEA, n.d.).

## Appendix C. Attrition Rates for Outcome Measures

Exhibits C1 through C9 list the overall attrition (i.e., the rate of attrition for the entire sample) and differential attrition (i.e., the difference in the rates of attrition for the BARR and control groups) for the student outcome measures for the combined sample. The combinations of overall and differential attrition for confirmatory academic outcomes (Northwest Evaluation Association [NWEA] scores and core failure) result in low levels of potential bias even under the conservative assumptions of the What Works Clearinghouse (WWC) standards.<sup>16</sup> The combinations for the student survey outcome measure result in potentially acceptable levels of bias depending on which assumptions are used: The attrition was high under conservative assumptions but low under liberal assumptions of the WWC standards.

### Exhibit C1. Attrition for NWEA Reading Scores (10 schools)

Data Source	BARR	Control	Total
Assigned sample	1,667	2,271	3,938
Analytic sample	1,147	1,448	2,595
Overall attrition rate	31.2%	36.2%	34.1%
Differential rate			-5.1%

### Exhibit C2. Attrition for NWEA Mathematics Scores (10 schools)

Data Source	BARR	Control	Total
Assigned sample	1,667	2,271	3,938
Analytic sample	1,173	1,542	2,715
Overall attrition rate	29.6%	32.1%	31.1%
Differential rate			-2.5%

### Exhibit C3. Attrition for Core Course Failure (11 schools)

Data Source	BARR	Control	Total
Assigned sample	1,785	2,383	4,168
Analytic sample	1,467	1,916	3,383
Overall attrition rate	17.8%	19.6%	18.8%
Differential rate			1.8%

<sup>16</sup> According to WWC, the choice of liberal or conservative assumptions is based on the relationship between treatment status and attrition. When attrition is not related to treatment status, liberal assumptions may be appropriate, when attrition is related to treatment status, conservative assumptions may be appropriate.



**Exhibit C4. Attrition for Supportive Relationships (11 schools)**

Data Source	BARR	Control	Total
Assigned sample	1785	2383	4168
Analytic sample	1251	1463	2714
Overall attrition rate	29.9%	38.6%	34.9%
Differential rate			-8.69%

**Exhibit C5. Attrition for Expectations and Rigor (11 schools)**

Data Source	BARR	Control	Total
Assigned sample	1785	2383	4168
Analytic sample	1263	1480	2743
Overall attrition rate	29.2%	37.9%	34.2%
Differential rate			-8.6%

**Exhibit C6. Attrition for Student Engagement (11 schools)**

Data Source	BARR	Control	Total
Assigned sample	1785	2383	4168
Analytic sample	1253	1466	2719
Overall attrition rate	29.8%	38.5%	34.8%
Differential rate			-8.7%

**Exhibit C7. Attrition for Sense of Belonging (11 schools)**

Data Source	BARR	Control	Total
Assigned sample	1785	2383	4168
Analytic sample	1243	1447	2690
Overall attrition rate	30.4%	39.3%	35.5%
Differential rate			-8.9%

**Exhibit C8. Attrition for Social and Emotional Learning (11 schools)**

Data Source	BARR	Control	Total
Assigned sample	1785	2383	4168
Analytic sample	1243	1455	2698
Overall attrition rate	30.4%	38.9%	35.3%
Differential rate			-8.6%

**Exhibit C9. Attrition for Grit (11 schools)**

Data Source	BARR	Control	Total
Assigned sample	1785	2383	4168
Analytic sample	1239	1444	2683
Overall attrition rate	30.6%	39.4%	35.6%
Differential rate			-8.8%

## Appendix D. Analysis of Baseline Equivalence

Exhibits in this appendix show baseline characteristics for all analytic samples we used in the study. The last column in each table shows the standardized mean difference (SMD) between treatment and comparison groups of students.

**Exhibit D1. Treatment and Comparison Group Characteristics at Baseline for NWEA Reading (10 Schools)**

Baseline Covariates	Treatment (N = 1,147)	Comparison (N = 1,448)	SMD
Pretest (Mathematics)	224.9	225.2	0.02
Pretest (Reading)	216.0	216.6	0.04
Female (%)	50.0	50.0	0.00
2qsHispanic (%)	57.3	63.7	0.13
White (%)	26.7	20.4	0.15
Black (%)	8.9	8.6	0.01
English Learner (%)	29.1	35.2	0.13
Special Education (%)	7.7	7.5	0.01
Free or Reduced-Price Lunch (%)	75.4	77.6	0.05

**Exhibit D2. Treatment and Comparison Group Characteristics at Baseline for NWEA Mathematics Baseline Covariates (10 Schools)**

Baseline Covariates	Treatment (N = 1,173)	Comparison (N = 1,542)	SMD
Pretest (Mathematics)	224.8	224.8	0.00
Pretest (Reading)	215.6	216.0	0.02
Female (%)	49.4	52.8	0.01
Hispanic (%)	57.8	64.2	0.13
White (%)	25.7	19.5	0.15
Black (%)	9.3	9.1	0.01
English Learner (%)	29.8	36.3	0.14
Special Education (%)	7.7	7.1	0.02
Free or Reduced-Price Lunch (%)	75.9	78.3	0.06

**Exhibit D3. Treatment and Comparison Group Characteristics at Baseline for Course Failure (11 Schools)**

Baseline Covariates	Treatment (N = 1,467)	Comparison (N = 1,916)	SMD
Pretest (Mathematics)	224.4	17.8	0.06
Pretest (Reading)	215.6	16.8	0.02
Female (%)	48.4	48.7	0.01
Hispanic (%)	53.3	59.7	0.13
White (%)	30.1	23.2	0.16
Black (%)	9.6	10.1	0.02
English Learner (%)	28.4	34.2	0.13
Special Education (%)	7.8	7.8	0.00
Free or Reduced-Price Lunch (%)	77.5	79.3	0.05

**Exhibit D4. Treatment and Comparison Group Characteristics at Baseline for Supportive Relationships (11 Schools)**

Baseline Covariates	Treatment (N = 1,251)	Comparison (N = 1,463)	SMD
Pretest (Mathematics)	225.3	225.5	0.01
Pretest (Reading)	216.5	216.8	0.02
Female (%)	49.1	49.8	0.02
Hispanic (%)	57.6	63.4	0.12
White (%)	29.8	23.2	0.15
Black (%)	8.2	8.5	0.01
English Learner (%)	28.7	34.4	0.12
Special Education (%)	6.6	5.8	0.03
Free or Reduced-Price Lunch (%)	76.3	78.1	0.04

**Exhibit D5. Treatment and Comparison Group Characteristics at Baseline for Expectation and Rigor (11 Schools)**

Baseline Covariates	Treatment (N = 1,263)	Comparison (N = 1,480)	SMD
Pretest (Mathematics)	225.2	225.5	0.02
Pretest (Reading)	216.3	216.9	0.03
Female (%)	49.1	49.6	0.01
Hispanic (%)	57.5	63.2	0.12
White (%)	30.0	23.4	0.15
Black (%)	8.2	8.4	0.01
English Learner (%)	28.7	34.5	0.13
Special Education (%)	6.7	5.8	0.04
Free or Reduced-Price Lunch (%)	76.4	78.0	0.04

**Exhibit D6. Treatment and Comparison Group Characteristics at Baseline for Student Engagement (11 Schools)**

Baseline Covariates	Treatment (N = 1243)	Comparison (N = 1,447)	SMD
Pretest (Mathematics)	225.3	225.6	0.02
Pretest (Reading)	216.5	217.0	0.03
Female (%)	49.2	50.2	0.02
Hispanic (%)	57.4	63.1	0.12
White (%)	30.0	23.3	0.15
Black (%)	8.2	8.6	0.01
English Learner (%)	28.5	34.3	0.13
Special Education (%)	6.6	5.9	0.03
Free or Reduced-Price Lunch (%)	76.5	77.9	0.03

**Exhibit D7. Treatment and Comparison Group Characteristics at Baseline for Sense of Belonging (11 Schools)**

Baseline Covariates	Treatment (N = 1,253)	Comparison (N = 1,466)	SMD
Pretest (Mathematics)	225.3	225.5	0.01
Pretest (Reading)	216.5	216.9	0.02
Female (%)	49.2	49.7	0.01
Hispanic (%)	57.4	63.3	0.12
White (%)	30.0	23.3	0.15
Black (%)	8.2	8.5	0.01
English Learner (%)	28.7	34.4	0.13
Special Education (%)	6.6	5.8	0.03
Free or Reduced-Price Lunch (%)	76.2	78.0	0.04

**Exhibit D8. Treatment and Comparison Group Characteristics at Baseline for Social and Emotional Learning (11 Schools)**

Baseline Covariates	Treatment (N = 1,243)	Comparison (N = 1,455)	SMD
Pretest (Mathematics)	225.3	225.5	0.01
Pretest (Reading)	216.5	216.8	0.02
Female (%)	49.2	50.0	0.02
Hispanic (%)	57.5	63.2	0.12
White (%)	29.9	23.3	0.15
Black (%)	8.2	8.5	0.01
English Learner (%)	28.7	34.4	0.12
Special Education (%)	6.5	5.8	0.03
Free or Reduced-Price Lunch (%)	76.4	77.9	0.04

**Exhibit D9. Treatment and Comparison Group Characteristics at Baseline for Grit (11 Schools)**

Baseline Covariates	Treatment (N = 1,239)	Comparison (N = 1,444)	SMD
Pretest (Mathematics)	225.3	225.5	0.01
Pretest (Reading)	216.5	216.9	0.02
Female (%)	49.3	50.1	0.02
Hispanic (%)	57.4	63.2	0.12
White (%)	30.0	23.2	0.16
Black (%)	8.2	8.6	0.02
English Learner (%)	28.5	34.3	0.13
Special Education (%)	6.6	5.9	0.03
Free or Reduced-Price Lunch (%)	76.6	78.1	0.04

## Appendix E. 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). 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 program 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 strategies. Two evaluators observed and rated three BARR activities during spring site visits each year (2015–17) using fidelity rubrics designed in advance of data collection (See Appendix B). Exhibit E1 provides the number of observations of each activity collected for each participating school in the study.

**Exhibit E1. Observations Collected From Participating Schools**

School Name	Cohort	Block/Team Meetings	I-Time Lessons	Risk Reviews
School A	1	1	1	1
School B	1	3	3	1
School C	1	2	2	1
School D	2	1	2	1
School E	2	2	2	1
School F	2	1	2	1
School G	3	1	2	1
School H	3	1	2	1
School I	3	1	2	1
School J	3	2	2	1
School K	3	2	2	1

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 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 nearly completely to the extent intended.

After the structural component review interview, we combined BARR coordinators' ratings with observation data. For the three observed BARR activities (block/team meetings, I-Time lessons, and risk-review meetings), the final component fidelity rating is based on an equal weight of the rating from the BARR coordinator (50 percent) and from the average score of observations made during the site visit (50 percent). All other fidelity ratings are based on interviews with the BARR coordinators.

### Setting Thresholds for Measurement

Our assessment of implementation fidelity occurred annually, in the summer, after the first year of BARR implementation in each school. These assessments were based on predetermined thresholds for adequacy of program implementation for each of the BARR key components or strategies (e.g., for the professional development indicator, a school would have had to score a 4.7 or higher to be considered as implementing with adequate fidelity). The fidelity thresholds for indicators under each key component were determined on the basis of theoretical expectations and practical experiences at the beginning of the study, in consultation with the program developers.

After Cohort 1 data collection and analyses were complete, we discussed revisions to the BARR coordinator review form with BARR developers in February 2016. Several indicators were added, and the scoring rubric was revised with detailed information for each implementation category. The changes to these indicators were made to reflect elements that we and the developers felt may have been missed in the previous version, and the updated version was

subsequently used for Cohorts 2 and 3. The fidelity matrix is provided in Exhibit E2. For more information of the indicators under each component, please contact BARR directly.

**Exhibit E2. Fidelity Matrix**

Component	Data Source	Indicator-Level Metric	Threshold for Adequate Implementation for a School	Threshold for Adequate Implementation for Cohort and Full Sample
Whole Student Emphasis	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if <i>Not yet present</i> , 3–5 if <i>Emerging</i> , 6–7 if <i>In Place</i>	Mean of four indicators at or above 5.5 for schools in Cohorts 1, 2, and 3	Two-thirds (67%) of schools rated as having adequate school-level implementation on this component
Professional Development	Hazelden training evaluations, BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if <i>Not yet present</i> , 3–5 if <i>Emerging</i> , 6–7 if <i>In Place</i>	Mean of three indicators at or above 4.7 for schools in Cohorts 1, 2, and 3	Two-thirds (67%) of schools rated as having adequate school-level implementation on this component
I-Time	BARR coordinator interview, observation data	Up to 7 points total for each indicator; 1–2 if <i>Not yet present</i> , 3–5 if <i>Emerging</i> , 6–7 if <i>In Place</i>	Mean of three indicators at or above 6.0 for schools in Cohort 1 and mean of six indicators at or above 5.3 for schools in Cohorts 2 and 3	Two-thirds (67%) of schools rated as having adequate school-level implementation on this component
Restructuring	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if <i>Not yet present</i> , 3–5 if <i>Emerging</i> , 6–7 if <i>In Place</i>	Mean of six indicators at or above 5.3 for schools in Cohort 1 and mean of five indicators at or above 5.3 for schools in Cohorts 2 and 3	Two-thirds (67%) of schools rated as having adequate school-level implementation on this component

Component	Data Source	Indicator-Level Metric	Threshold for Adequate Implementation for a School	Threshold for Adequate Implementation for Cohort and Full Sample
Block/Team Meeting	BARR coordinator interview, observation data	Up to 7 points total for each indicator; 1–2 if <i>Not yet present</i> , 3–5 if <i>Emerging</i> , 6–7 if <i>In Place</i>	Mean of five indicators at or above 5.6 for schools in Cohort 1 and mean of seven indicators at or above 5.6 for schools in Cohorts 2 and 3	Two-thirds (67%) of schools rated as having adequate school-level implementation on this component
Risk Review	BARR coordinator interview, observation data	Up to 7 points total for each indicator; 1–2 if <i>Not yet present</i> , 3–5 if <i>Emerging</i> , 6–7 if <i>In Place</i>	Mean of five indicators at or above 4.8 for schools in Cohort 1 and mean of six indicators at or above 4.8 for schools in Cohorts 2 and 3	Two-thirds (67%) of schools rated as having adequate school-level implementation on this component
Parent Involvement	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if <i>Not yet present</i> , 3–5 if <i>Emerging</i> , 6–7 if <i>In Place</i>	Mean of six indicators at or above 3.3 for schools in Cohorts 1, 2, and 3	Two-thirds (67%) of schools rated as having adequate school-level implementation on this component
Contextual Supports	BARR coordinator interview	Up to 7 points total for each indicator; 1–2 if <i>Not yet present</i> , 3–5 if <i>Emerging</i> , 6–7 if <i>In Place</i>	Mean of seven indicators at or above 4.0 for schools in Cohorts 1, 2, and 3	Two-thirds (67%) of schools rated as having adequate school-level implementation on this component

Exhibit E3 provides a summary of the number of indicators for each BARR strategy that were used for the Cohort 1 schools and in the updated version. Thresholds for adequate implementation remained the same for all components except for I-Time. For this component, the number of indicators increased from three to six, which changed the threshold for adequate implementation from 6.0 to 5.3 on the basis of different expectations for teachers for those new indicators. Note that changes between cohorts are highlighted in light blue.

**Exhibit E3. Changes in Number of Indicators and Thresholds From Participating Schools**

BARR Key Component/ Strategies on Logic Model	Cohort 1 (2014–15)		Cohorts 2 and 3 (2015–17)	
	Indicators	Threshold	Indicators	Threshold
Professional Development	3	4.7 or higher	3	4.7 or higher
Restructuring Ninth Grade	6	5.3 or higher	5	5.3 or higher
Whole Student Emphasis	4	5.5 or higher	4	5.5 or higher
Block/Team Meetings	5	5.6 or higher	7	5.6 or higher
I-Time	3	6.0 or higher	6	5.3 or higher
Risk Review	5	4.8 or higher	6	4.8 or higher
Contextual Support	7	4.0 or higher	7	4.0 or higher
Parent Involvement	6	3.3 or higher	6	3.3 or higher

## Appendix F. Statistical Models

Appendix F describes the analytic models used to estimate program effects on student and teacher outcomes and details the total number of imputed cases for the three main impact estimates.

### F1. Student Outcomes Model

We estimated program effects on all student outcomes (i.e., Northwest Evaluation Association [NWEA] Measures of Academic Progress (MAP) Reading scores, NWEA MAP Mathematics scores, core course failure student survey measures) presented in this report by using an ordinary least squares model to compare outcomes for students assigned to the BARR group with outcomes for students assigned to the control group. All impact models included student-level background characteristics (e.g., race, gender, FRPL status, ELL status), a test of prior student achievement (i.e., students' fall NWEA MAP scores), an indicator of a student's assignment to BARR, and a set of dummy variables to control for school effects and variation in the random-assignment ratio across schools. Of note, we imputed missing baseline variables and pretest scores using the dummy variable imputation methods recommended by Puma and colleagues (2009). See Appendix F3.

Analytic Model

$$Y_i = \beta_0 + \beta_1 \text{BARR}_i + \mathbf{B}_2 \mathbf{X}_i + \mathbf{B}_3 \mathbf{Z}_j + \varepsilon_i$$

$Y_i$  is the outcome for individual  $i$ ;

$\beta_0$  is the mean outcome for the control group;

**$\beta_1$  is the true program effect;**

$\text{BARR}_i$  is an indicator variable equal to 1 for BARR group members and 0 for control group members;

**$\mathbf{B}_2$**  is a vector of student level predictors;

**$\mathbf{X}$**  is a vector of student characteristics including gender, FRL, LEP, special education, and pretest scores;

**$\mathbf{B}_3$**  is a vector of indicator variables representing each school  $j$ ;

**$\mathbf{Z}$**  is a series of school indicators; and

$\varepsilon_i$  = the error component for individual  $i$ .

## F2. Teacher Outcomes Model

An independent-sample *t*-test was conducted to compare BARR teachers with control teachers in the experiences measured by the eight scales on the teacher survey.

## F3. Imputed Baseline Variables

**Exhibit F3.1. The Total Number (%) of Imputed Missing Cases for NWEA Reading (10 Schools)**

Baseline Variable	BARR (N = 1,147)	Comparison (N = 1,448)	Total (N = 2,595)
Pretest (Mathematics) missing (%)	219 (19.1%)	226 (15.6%)	445 (17.2%)
Pretest (Reading) missing (%)	178 (15.5%)	199 (13.7%)	377 (14.5%)
Free or Reduced-Price Lunch missing (%)	5 (0.4%)	6 (0.4%)	11 (0.4%)

**Exhibit F3.2. The Total Number (%) of Imputed Missing Cases for NWEA Mathematics (10 Schools)**

Baseline Variable	BARR (N = 1173)	Comparison (N = 1542)	Total (N = 2715)
Pretest (Mathematics) missing (%)	229 (19.5%)	234 (15.2%)	463 (17.1%)
Pretest (Reading) missing (%)	184 (15.7%)	212 (13.7%)	396 (14.6%)
Free or Reduced-Price Lunch missing (%)	12 (1.0%)	0 (0.0%)	12 (0.4%)

**Exhibit F3.3. The Total Number (%) of Imputed Missing Cases for Course Failure (11 Schools)**

Baseline Variable	BARR (N = 1,467)	Comparison (N = 1,916)	Total (N = 3,383)
Pretest (Mathematics) missing (%)	287 (19.6%)	314 (16.4%)	601 (17.8%)
Pretest (Reading) missing (%)	239 (16.3%)	289 (15.1%)	528 (15.6%)
Free or Reduced-Price Lunch missing (%)	7 (0.5%)	14 (0.7%)	21 (0.6%)

## Appendix G. NWEA Sensitivity Analyses

This section details the impact of BARR, after one year of implementation, on student Northwest Evaluation Association (NWEA) reading and mathematics scores and growth projections for five study school in which (a) we had valid scores on NWEA assessments for at least 65 percent of the students, and (b) we had differential attrition lower than 5 percentage points across the research groups. These sensitivity analyses did not find meaningfully different impact results than we estimated for the ten schools.

**Exhibit G1. Impacts on Standardized Reading Scale Scores by Subgroup (Five schools)**

Outcome/Subgroup	N	BARR	Control	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>1,373</b>	<b>221.94</b>	<b>222.42</b>	<b>-0.48</b>	<b>0.308</b>	<b>0.03</b>
Female	719	223.52	223.94	-0.43	0.447	0.03
Male	654	220.29	220.68	-0.40	0.612	0.03
Minority	988	219.50	220.17	-0.67	0.254	0.05
White	385	227.97	228.28	-0.30	0.699	0.02
Free or Reduced-Price Lunch	980	220.00	220.15	-0.15	0.796	0.01
Not Free or Reduced-Price Lunch	391	226.91	228.18	-1.27	0.126	0.09
English Learners	471	218.11	219.84	-1.72*	0.045	0.12
Non-English Learners	902	223.75	223.80	-0.05	0.928	0.00
Special Education	108	205.82	209.05	-3.24	0.117	0.20
Non-Special Education	1,265	223.28	223.59	-0.31	0.529	0.02

Source: AIR calculations from school-administered NWEA assessments.

Note. \* = statistically significant at the  $p < .05$  level.

**Exhibit G2. Impacts on Standardized Mathematics Scale Scores by Subgroup (Five schools)**

Outcome/Subgroup	N	BARR	Control	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>1,403</b>	<b>232.43</b>	<b>232.00</b>	<b>0.43</b>	<b>0.349</b>	<b>0.02</b>
Female	726	231.42	231.69	-0.27	0.638	0.02
Male	677	233.39	232.39	1.00	0.186	0.05
Minority	1,008	228.43	228.57	-0.14	0.808	0.01

Outcome/Subgroup	N	BARR	Control	Difference	P-Value	Effect Size
White	395	242.56	240.45	2.11†	0.007	0.12
Free or Reduced-Price lunch	1,002	229.32	228.92	0.40	0.492	0.02
Not Free or Reduced-Price Lunch	399	240.28	239.84	0.44	0.568	0.02
English Learners	481	227.00	228.81	-1.81*	0.017	0.10
Non-English Learners	922	235.08	233.63	1.46*	0.012	0.08
Special Education	111	215.20	214.07	1.13	0.553	0.06
Non-Special Education	1,292	233.91	233.53	0.38	0.424	0.02

Source: AIR calculations from school-administered NWEA assessments.

Note. \* = statistically significant at the  $p < .05$  level; † = statistically significant at the  $p < .01$  level.

**Exhibit G3. Impacts on Reading Growth Projections Met by Subgroup (Five schools)**

Outcome/Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
<b>Full Sample</b>	<b>1,348</b>	<b>58.7</b>	<b>57.1</b>	<b>1.6</b>	<b>0.565</b>	<b>0.04</b>
Female	708	58.9	58.7	0.2	0.963	0.00
Male	640	58.1	55.7	2.4	0.552	0.06
Minority	966	57.3	55.7	1.6	0.643	0.04
White	382	62.1	60.9	1.2	0.805	0.03
Free or Reduced-Price Lunch	958	57.8	56.1	1.7	0.621	0.04
Not Free or Reduced-Price Lunch	388	60.3	59.8	0.5	0.924	0.01
English Learners	460	53.5	57.4	-3.9	0.413	0.10
Non-English Learners	888	60.6	57.1	3.5	0.298	0.09
Special Education	105	45.1	59.1	-14.0	0.174	0.36
Non-Special Education	1,243	59.8	57	2.8	0.331	0.07

Source: AIR calculations from school-administered NWEA assessments.



**Exhibit G4. Impacts on Mathematics Growth Projections Met by Subgroup (Five schools)**

Outcome/Subgroup	N	BARR (%)	Control (%)	Difference	P-value	Effect Size
<b>Full Sample</b>	<b>1,374</b>	<b>65.5</b>	<b>63.8</b>	<b>1.7</b>	<b>0.521</b>	<b>0.05</b>
Female	712	63.7	68.5	-4.8	0.182	0.13
Male	662	67.4	58.5	8.9*	0.021	0.24
Minority	981	62.6	65.1	-2.5	0.430	0.07
White	393	71.1	59.6	11.5*	0.018	0.32
Free or Reduced-Price lunch	978	63.8	63.0	0.8	0.803	0.02
Not Free or Reduced-Price lunch	394	69.3	66.0	3.3	0.484	0.09
English Learners	471	60.4	68.5	-8.1	0.076	0.22
Non-English Learners	903	67.4	61.1	6.3	0.053	0.17
Special Education	110	58.6	59.4	-0.8	0.933	0.02
Non-Special Education	1,264	66.1	64.2	1.9	0.495	0.05

Source: AIR calculations from school-administered NWEA assessments.

Note. \* = statistically significant at the  $p < .05$  level.

## Appendix H. Supplemental Student Behavior Subgroup Analyses

Appendix H provides the subgroup analyses for the impacts of BARR on three exploratory outcomes: the percentage of students who were chronically absent (defined as missing 10% or more days), the percentage of students who were ever suspended (in school or out of school) during the year, and the percentage of students who enrolled in 10th grade in the same district the following year.

### Exhibit H1. Impacts on Students’ Chronic Absenteeism by Subgroup (Nine schools)

Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
<b>Full Sample</b>	3,275	23.1	21.9	1.2	0.314	0.04
Female	1,537	23.5	21.2	2.2	0.214	0.08
Male	1,636	22.9	22.5	0.5	0.780	0.02
Minority	2,230	21.6	21.8	-0.2	0.892	0.01
White	942	26.0	22.2	3.7	0.106	0.12
Free or Reduced-Price Lunch	2,330	21.3	20.6	0.8	0.599	0.03
Not Free or Reduced-Price Lunch	723	18.0	17.0	1.0	0.658	0.04
English Learners	875	16.8	15.7	1.2	0.576	0.05
Non-English Learners	2,400	28.9	27.4	1.5	0.319	0.04
Special Education	243	27.9	31.1	-3.2	0.514	0.09
Non-Special Education	3,032	25.4	23.7	1.7	0.183	0.06

Source: AIR calculations from school-provided administrative data.

### Exhibit H2. Impacts on Students’ Suspensions by Subgroup (10 schools)

Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
<b>Full Sample</b>	3,806	6.5	6.7	-0.2	0.774	0.02
Female	1,775	5.8	4.5	1.4	0.250	0.17
Male	1,929	8.2	9.7	-1.5	0.258	0.11
Minority	2,743	6.1	5.5	0.6	0.536	0.06
White	960	8.2	10.5	-2.3	0.192	0.17
Free or Reduced-Price Lunch	2,825	6.9	7.3	-0.4	0.679	0.04

Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Not Free or Reduced-Price Lunch	759	5.9	4.6	1.3	0.467	0.15
English Learners	1,132	5.1	3.4	1.7	0.201	0.26
Non-English Learners	2,674	6.8	7.7	-0.9	0.359	0.08
Special Education	294	11.4	19.0	-7.6	0.053	0.36
Non-Special Education	3,512	5.8	5.5	0.4	0.638	0.04

Source: AIR calculations from school-provided administrative data.

**Exhibit H3. Impacts on Students’ Persistence to Grade 10 by Subgroup (11 schools)**

Subgroup	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
<b>Full Sample</b>	2,863	88.8	87.4	1.4	0.222	0.08
Female	1,408	87.6	88.4	-0.8	0.618	0.05
Male	1,455	89.6	85.6	4.0*	0.024	0.22
Minority	1,996	86.8	85.8	1.0	0.523	0.05
White	867	93.4	90.4	3.1	0.083	0.25
Free or Reduced-Price Lunch	2,167	88.2	86.6	1.5	0.273	0.09
Not Free or Reduced-Price Lunch	664	94.3	90.9	3.4	0.087	0.31
English Learners	807	91.2	92.1	-0.9	0.638	0.07
Non-English Learners	2,056	87.9	85.5	2.4	0.105	0.12
Special Education	214	86.5	84.0	2.5	0.613	0.12
Non-Special Education	2,649	88.9	87.8	1.1	0.345	0.07

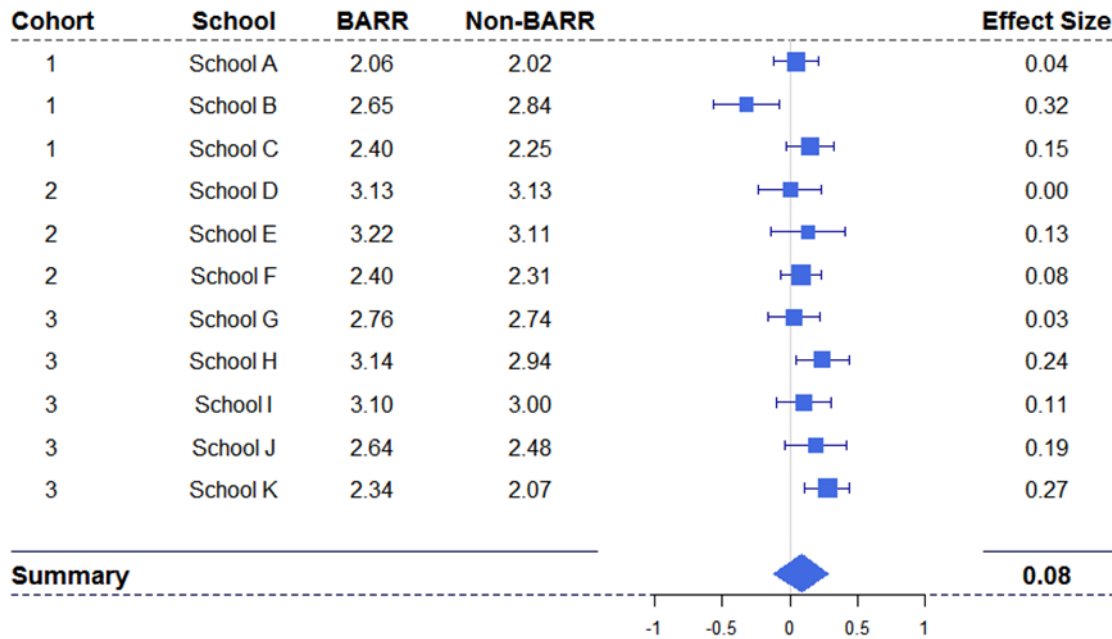
Source: AIR calculations from school-provided administrative data.

Note. \* = statistically significant at the  $p < .05$  level.

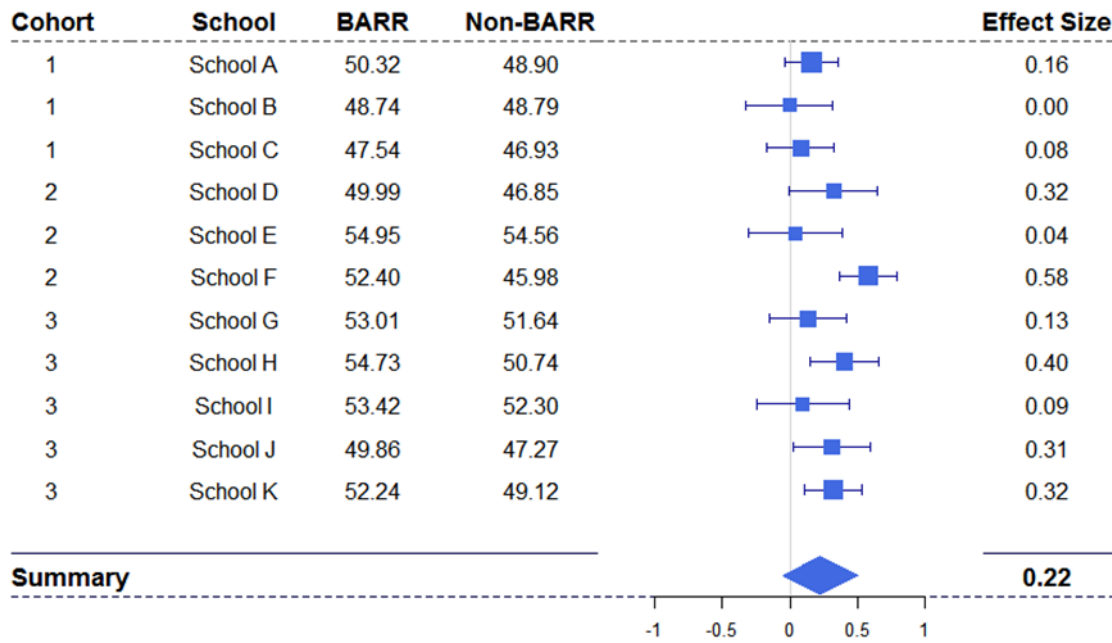
## Appendix I. Supplemental Impact Estimates by School

Appendix I includes estimates of treatment effects by school for grade point average, student experiences, and behavior outcomes.

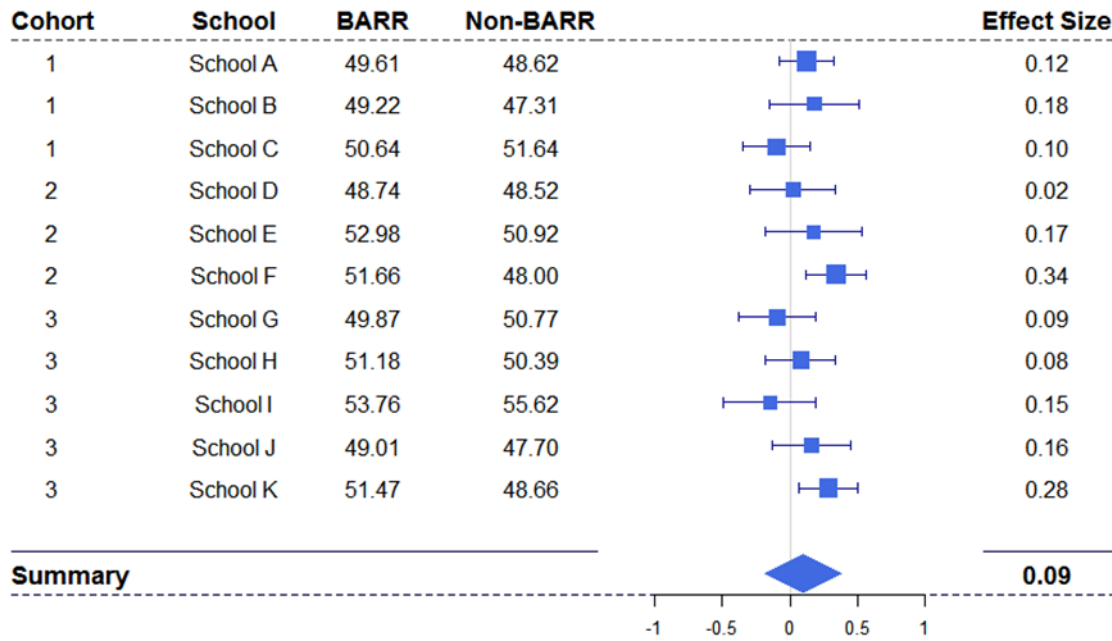
**Exhibit I1. Estimated Effects on Students’ Grade Point Average, by School (11 Schools)**



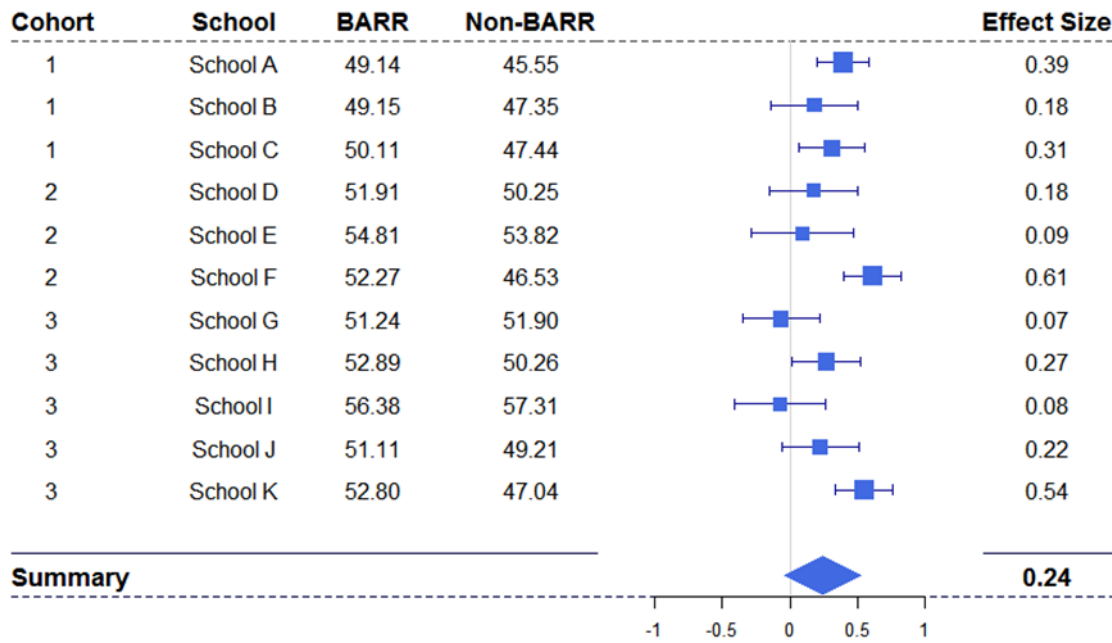
**Exhibit I2. Estimated Effects on Students’ Expectations and Rigor, by School (11 schools)**



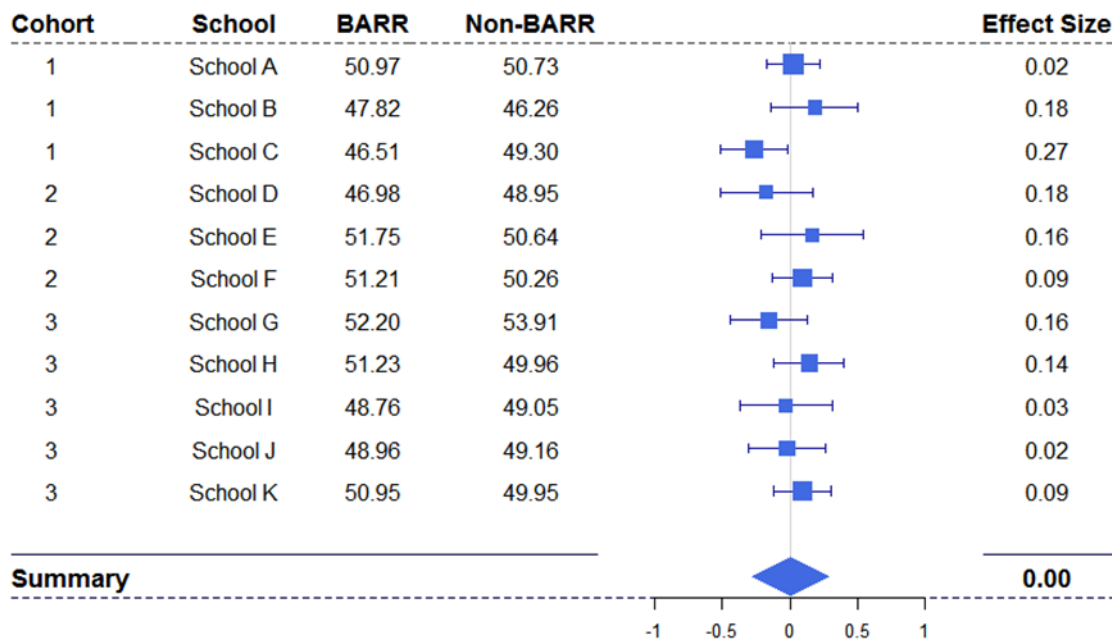
**Exhibit I3. Estimated Effects on Students’ Engagement, by School (11 schools)**



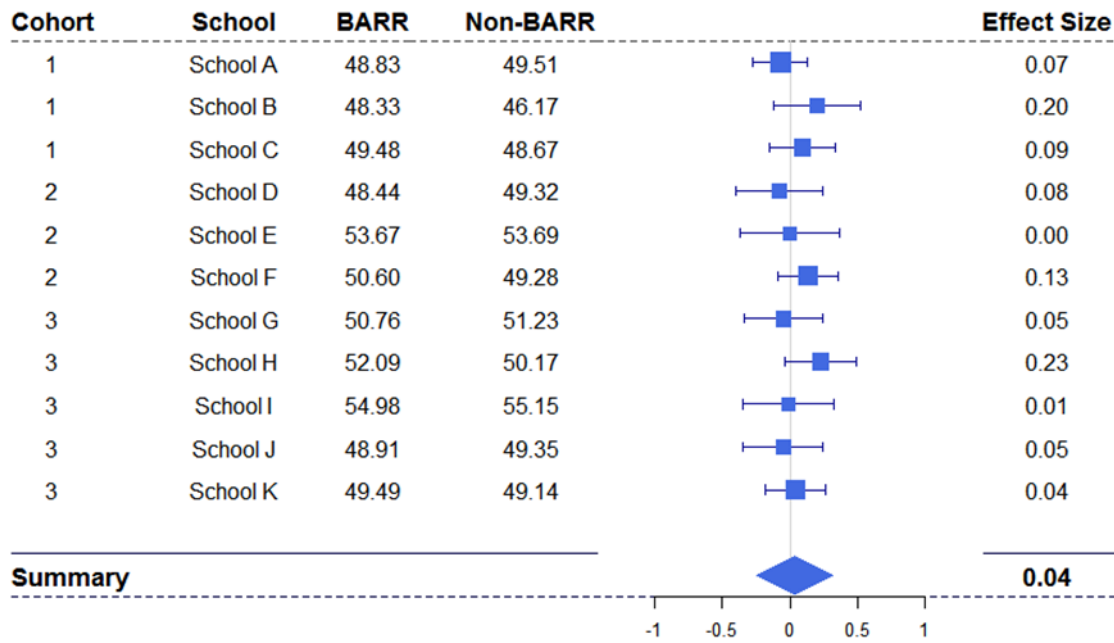
**Exhibit I4. Estimated Effects on Students’ Supportive Relationships, by School (11 schools)**



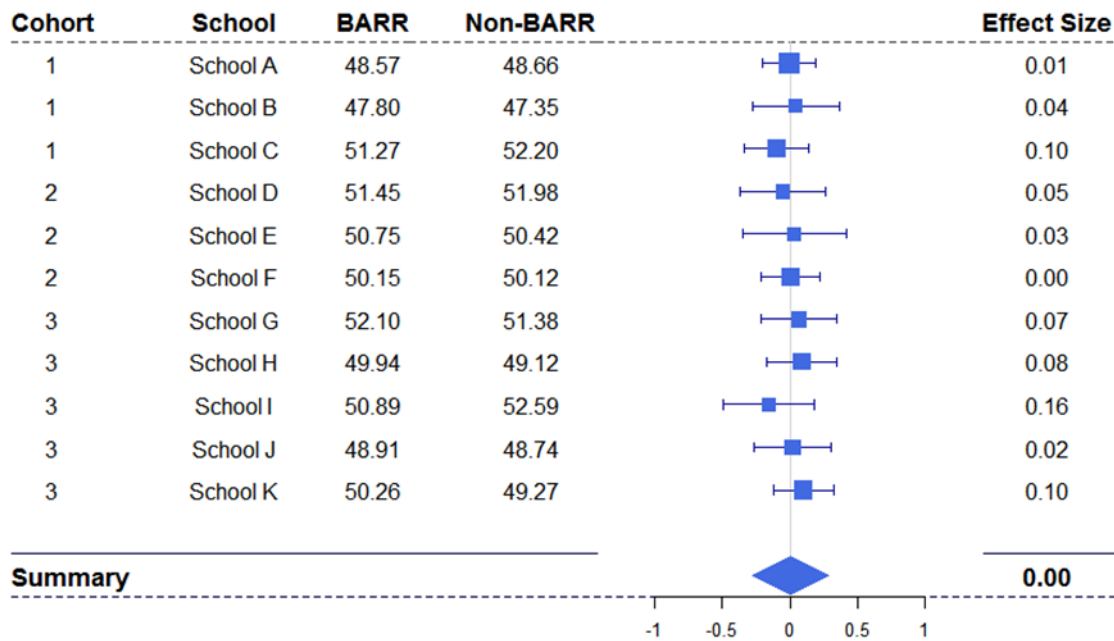
**Exhibit I5. Estimated Effects on Students’ Social and Emotional Learning, by School (11 Schools)**



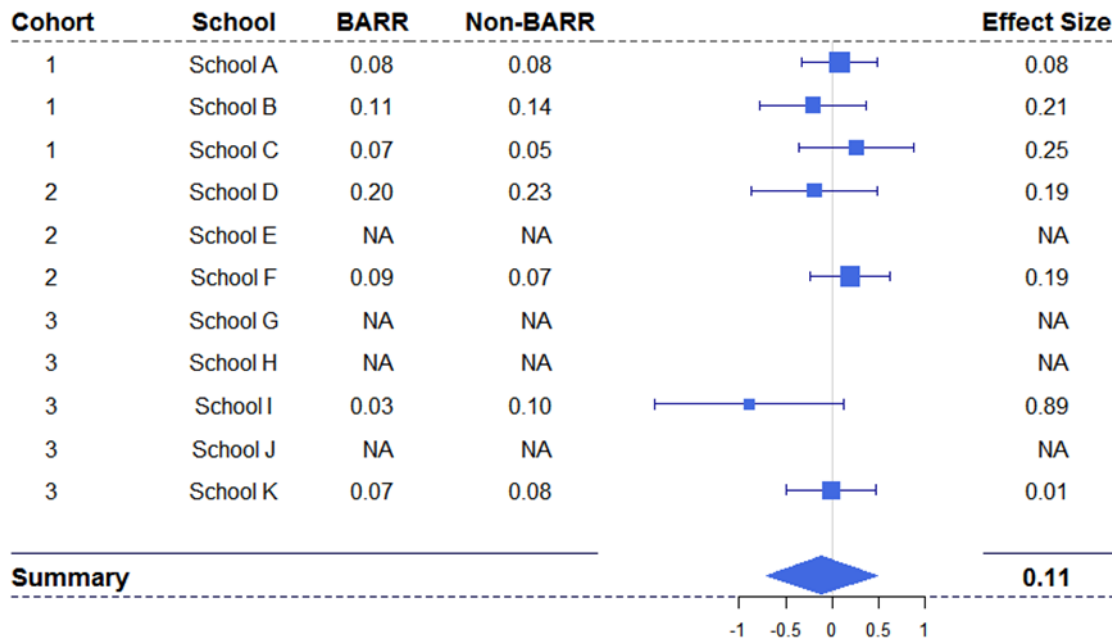
**Exhibit I6. Estimated Effects on Students’ Sense of Belonging, by School (11 Schools)**



**Exhibit I7. Estimated Effects on Students’ Grit, by School (11 Schools)**

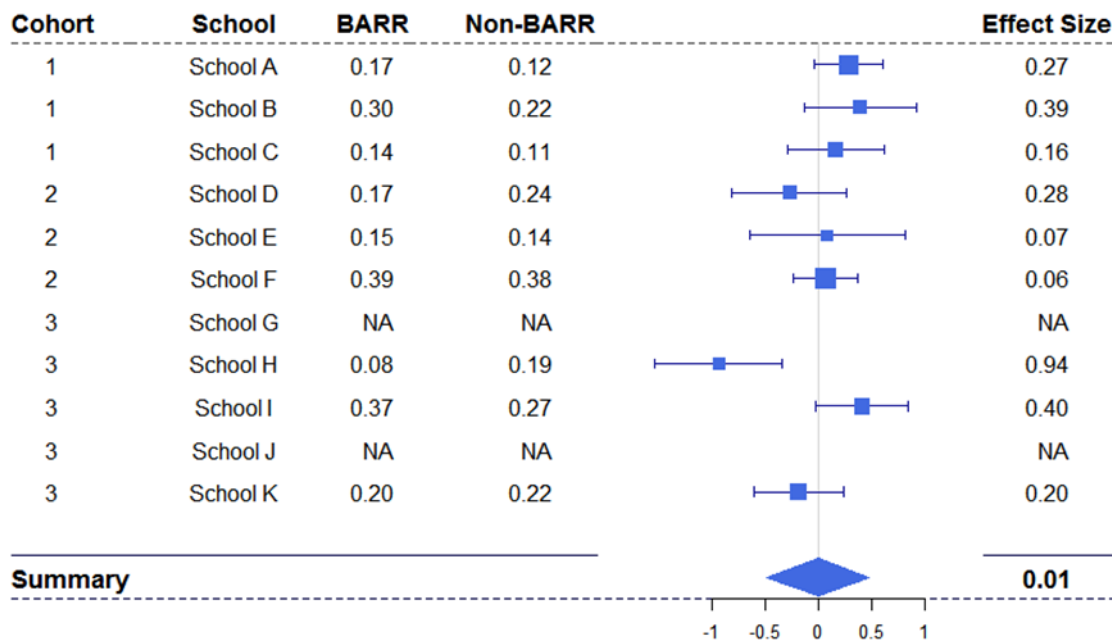


**Exhibit I8. Estimated Effects on Students’ Suspensions, by School (Seven Schools)**



Note. Reduction in suspensions is considered a favorable outcome for students.

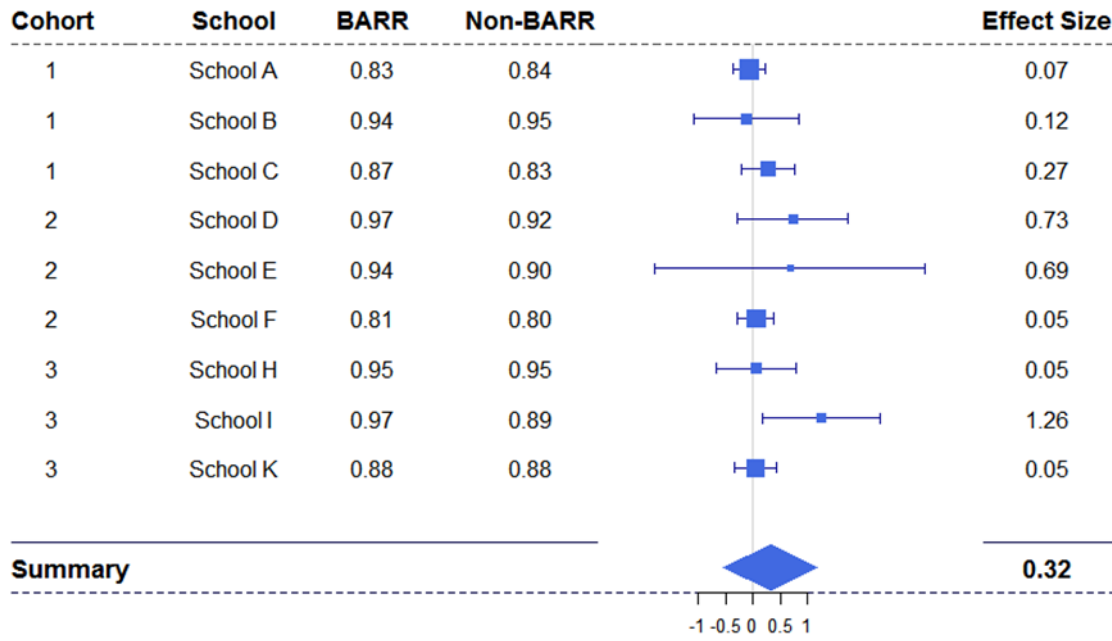
**Exhibit I9. Estimated Effects on Students’ Chronic Absenteeism, by School (Nine Schools)**



Note. Reduction in chronic absenteeism is considered a favorable outcome for students.



**Exhibit I10. Estimated Effects on Students’ Persistence to Grade 10, by School (11 Schools)**



## Appendix J. Supplemental Exploratory Impact Estimates by NWEA Baseline Test Score Quartile

This appendix provides exploratory impact estimates for student achievement, student experiences, and behavior outcomes broken down by the Northwest Evaluation Association (NWEA) baseline test score quartile.

### Exhibit J1. Impacts on Students Failing One or More Core Courses, by Quartile (11 Schools)

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	659	12.0	14.8	-2.79	0.336	0.15
Second Quartile	658	21.6	27.0	-5.46	0.098	0.18
Third Quartile	660	30.6	44.1	-13.41‡	0.000	0.35
Bottom Quartile	659	51.4	64.0	-12.65‡	0.001	0.32

Source: AIR calculations from school-provided administrative data

Note. ‡ = statistically significant at the  $p < .001$  level.

### Exhibit J2. Impacts on Students' Grade Point Average in Core Courses, by Quartile (11 Schools)

Quartile	N	BARR	Control	Difference	P-Value	Effect Size
Top Quartile	658	3.26	3.22	0.04	0.452	0.05
Second Quartile	658	2.78	2.73	0.05	0.418	0.06
Third Quartile	658	2.53	2.38	0.16*	0.010	0.18
Bottom Quartile	659	2.06	1.87	0.19†	0.002	0.21

Source: AIR calculations from school-provided administrative data .

Note. \* = statistically significant at the  $p < .05$  level; † = statistically significant at the  $p < .01$  level.

### Exhibit J3. Impacts on Standardized Reading Scale Scores, by Quartile (10 schools)

Quartile	N	BARR	Control	Difference	P-Value	Effect Size
Top Quartile	509	235.15	235.92	-0.77	0.241	0.08
Second Quartile	509	224.73	223.11	1.62*	0.016	0.20
Third Quartile	508	216.51	216.55	-0.03	0.968	0.00
Bottom Quartile	511	204.32	205.26	-0.93	0.338	0.07

Source: AIR calculations from school-administered NWEA assessments.

Note. \* = statistically significant at the  $p < .05$  level.

**Exhibit J4. Impacts on Standardized Mathematics Scale Scores, by Quartile (10 schools)**

Quartile	N	BARR	Control	Difference	P-Value	Effect Size
Top Quartile	529	249.53	249.39	0.14	0.813	0.01
Second Quartile	535	233.53	233.21	0.32	0.622	0.04
Third Quartile	534	224.00	222.96	1.04	0.156	0.10
Bottom Quartile	533	210.27	210.40	-0.13	0.892	0.01

Source: AIR calculations from school-administered NWEA assessments.

**Exhibit J5. Impacts on Reading Growth Projections Met, by Quartile (10 schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	515	62.4	62.3	0.1	0.969	0.00
Second Quartile	536	61.0	54.4	6.6	0.130	0.17
Third Quartile	491	52.0	43.6	8.4	0.068	0.22
Bottom Quartile	495	54.0	56.3	-2.3	0.613	0.06

Source: AIR calculations from school-administered NWEA assessments.

**Exhibit J6. Impacts on Mathematics Growth Projections Met, by Quartile (10 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	577	63.2	62.4	0.8	0.833	0.02
Second Quartile	565	58.8	56.9	1.9	0.635	0.06
Third Quartile	558	49.6	50.9	-1.3	0.768	0.03
Bottom Quartile	552	63.0	57.3	5.7	0.173	0.16

Source: AIR calculations from school-administered NWEA assessments.

**Exhibit J7. Impacts on Students’ Expectations and Rigor, by Quartile (11 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	541	51.70	50.00	1.70	0.051	0.17
Second Quartile	546	51.16	48.73	2.43†	0.007	0.23
Third Quartile	544	51.06	49.24	1.82*	0.030	0.19
Bottom Quartile	539	51.11	47.52	3.59‡	0.000	0.35

Source: AIR calculations from AIR-administered student survey.

Note. \* = statistically significant at the  $p < .05$  level; † = statistically significant at the  $p < .01$  level; ‡ = statistically significant at the  $p < .001$  level.

**Exhibit J8. Impacts on Students’ Engagement, by Quartile (11 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	539	50.46	50.59	-0.12	0.898	0.01
Second Quartile	542	50.53	49.67	0.86	0.326	0.09
Third Quartile	537	51.23	50.16	1.07	0.198	0.11
Bottom Quartile	533	50.36	48.56	1.81	0.062	0.17

Source: AIR calculations from AIR-administered student survey.

**Exhibit J9. Impacts on Students’ Supportive Relationships, by Quartile (11 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	536	52.12	50.53	1.58	0.076	0.15
Second Quartile	543	51.19	48.16	3.03‡	0.000	0.30
Third Quartile	538	52.19	49.51	2.68†	0.002	0.28
Bottom Quartile	532	52.12	48.53	3.53‡	0.000	0.34

Source: AIR calculations from AIR-administered student survey.

Note. † = statistically significant at the  $p < .01$  level; ‡ = statistically significant at the  $p < .001$  level.

**Exhibit J10. Impacts on Students’ Social and Emotional Learning, by Quartile (11 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	531	48.11	47.76	0.35	0.682	0.04
Second Quartile	543	48.38	48.77	-0.39	0.672	0.04
Third Quartile	537	51.16	51.16	0.00	0.998	0.00
Bottom Quartile	528	51.81	52.15	-0.34	0.690	0.04

Source: AIR calculations from AIR-administered student survey.

**Exhibit J11. Impacts on Students’ Sense of Belonging, by Quartile (11 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	531	51.25	50.25	1.00	0.273	0.10
Second Quartile	538	49.72	50.13	-0.42	0.644	0.04
Third Quartile	535	49.91	49.99	-0.08	0.927	0.01
Bottom Quartile	525	50.55	49.24	1.31	0.152	0.13

Source: AIR calculations from AIR-administered student survey.

**Exhibit J12. Impacts on Students’ Grit, by Quartile (11 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	527	51.59	52.11	-0.52	0.560	0.05
Second Quartile	538	49.79	51.41	-1.61	0.065	0.16
Third Quartile	535	50.52	49.71	0.82	0.335	0.08
Bottom Quartile	525	48.91	47.19	1.72	0.076	0.16

Source: AIR calculations from AIR-administered student survey.

**Exhibit J13. Impacts on Students’ Suspensions, by Subgroup (10 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	637	3.6	2.5	1.1	0.490	0.23
Second Quartile	632	8.6	6.8	1.8	0.438	0.15
Third Quartile	642	10.3	10.6	-0.4	0.902	0.02
Bottom Quartile	637	12.9	14.0	-1.1	0.707	0.06

Source: AIR calculations from school-provided administrative data.

**Exhibit J14. Impacts on Students’ Chronic Absenteeism, by Subgroup (Nine Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	557	7.3	8.8	-1.5	0.533	0.12
Second Quartile	556	12.4	9.5	2.9	0.298	0.18
Third Quartile	561	18.2	15.4	2.8	0.380	0.12
Bottom Quartile	558	16.4	18.5	-2.0	0.530	0.08

Source: AIR calculations from school-provided administrative data.

**Exhibit J15. Impacts on Students’ Persistence to Grade 10, by Subgroup (11 Schools)**

Quartile	N	BARR (%)	Control (%)	Difference	P-Value	Effect Size
Top Quartile	554	94.9	96.1	-1.2	0.504	0.17
Second Quartile	556	92.7	89.6	3.1	0.253	0.23
Third Quartile	557	87.1	91.1	-4.0	0.146	0.25
Bottom Quartile	556	87.8	81.2	6.6*	0.032	0.31

Source: AIR calculations from school-provided administrative data.

Note. \* = statistically significant at the  $p < .05$  level.



Established in 1946, the American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research on important social issues and delivers technical assistance, both domestically and internationally, in the areas of education, health, and workforce productivity.

## MAKING RESEARCH RELEVANT

AMERICAN INSTITUTES FOR RESEARCH  
1000 Thomas Jefferson Street NW  
Washington, DC 20007-3835 | 202.403.5000  
[www.air.org](http://www.air.org)

### LOCATIONS

**Domestic:** Washington, DC (HQ) | Monterey, Sacramento, and San Mateo, CA | Atlanta, GA | Honolulu, HI | Chicago and Naperville, IL | Indianapolis, IN | Metairie, LA | Waltham, MA | Frederick and Rockville, MD | Chapel Hill, NC | New York, NY | Columbus, OH | Cayce, SC | Austin, TX | Arlington and Reston, VA | Seattle, WA

**International:** Algeria | Ethiopia | Germany | Haiti | Zambia