Results

An Evaluation of the BellXcel Summer Program

Final Evaluation Report

June 2019

Prepared by:

Tracy J. Waters, Ph.D.
David Fleming, Ph.D.
Kelly Gregory, M.A.T.
Peyton Roth
Katherine Stevens
Furman University

Jennifer Laird, Ph.D.
Jessica Robles, Ed.D.
RTI International

The Riley Institute
FURMAN UNIVERSITY

This report is funded by the United Way of Greenville County's Social Innovation Fund grant to support its OnTrack Greenville partnership. The Social Innovation Fund (SIF) was a program that received funding from 2010 to 2016 from the Corporation for National and Community Service (CNCS), a federal agency that engages millions of Americans in service through its AmeriCorps, Senior Corps, and Volunteer Generation Fund programs, and leads the nation's volunteer and service efforts. Using public and private resources to find and grow community-based nonprofits with evidence of results, SIF intermediaries received funding to award subgrants that focus on overcoming challenges in economic opportunity, healthy futures, and youth development. Although CNCS made its last SIF intermediary awards in fiscal year 2016, SIF intermediaries will continue to administer their subgrant programs until their federal funding is exhausted.

United Way of Greenville County's OnTrack Greenville partnership includes support from CNCS and the following investors: Hollingsworth Funds, Community Foundation of Greenville, the Daniel-Mickel Foundation, Fluor Foundation, Gilreath Family Fund, the Graham Foundation, John I. Smith Charities, Jolley Foundation, Lockheed Martin, Piedmont Health Foundation, ScanSource Charitable Foundation, F.W. Symmes Foundation, Wells Fargo, JHM Hotels, Sisters of Charity Foundation of South Carolina, Priester Family Foundation, and SunTrust Foundation.

Suggested citation: Waters, T., Fleming, D., Laird, J., Robles, J., Gregory, K., Roth, P., & Stevens, K. (2019). *An Evaluation of the BellXcel Summer Program: Final Evaluation Report*. Greenville, SC: Furman University.

Leadership, Knowledge, Impact.

The Riley Institute at Furman broadens student and community perspectives about issues critical to South Carolina's progress. It builds and engages present and future leaders, creates and shares data-supported information about the state's core challenges, and links the leadership body to those data to drive sustainable solutions.

Launched in 1999, the Institute is named for former South Carolina governor and former United States Secretary of Education Richard W. (Dick) Riley. It is committed to nonpartisanship in all it does and to a rhetoric-free, facts-based approach to change.

The Riley Institute's research group promotes evidence-informed education practices and supports organizations that serve children and families across South Carolina. Part of the Institute's Center for Education Policy and Leadership, the group conducts in-depth research and evaluation studies to support organizational decision-making and practice. It also builds internal evaluation capacity among organizations serving children and families to maximize outcomes for citizens across the state.

For more information, please visit riley.furman.edu.

Table of Contents

List of Tables

List of Figures

Executive Summary

The United Way of Greenville County (UWGC) received an award in the 2014 Social Innovation Fund (SIF) grant competition to support its OnTrack Greenville initiative, a collective impact dropout-prevention program for middle grades students. BellXcel, a Sub-Grantee, implemented the BellXcel Summer Program, an evidence-based summer learning program that targets rising sixth graders at three of the four OnTrack Greenville schools. The Riley Institute at Furman University served as the third-party evaluation contractor for the SIF-funded evaluation of OnTrack Greenville, including the BellXcel Summer Program. BellXcel implemented its Summer Program at three middle schools in the White Horse Community of Greenville County, South Carolina.

In summer 2016 and 2017, BellXcel operated a rigorous and stimulating six-week summer program for approximately 240 rising sixth-grade students (80 at each school) who were identified as academically and/or behaviorally at-risk. For four days each week, enrolled students received three hours of academic instruction, three hours of enrichment courses, free transportation, a healthy breakfast and lunch, and the opportunity to engage in recreational activities and go on field trips. In addition, BellXcel offered a series of events for the students' families throughout the summer session in order to improve parent engagement. BellXcel's model is designed to create a seamless connection between the summer program and the academic year, strengthening the transition from elementary to middle school. The primary intended outcome of the BellXcel Summer Program was to improve student course performance in math and ELA. The secondary outcomes were to improve student self-confidence and attitude toward learning, among others.

BellXcel offered the Summer Program at three OnTrack Greenville treatment schools, aiming to serve approximately 240 middle school students each summer. In summer 2016, 233 students were enrolled in the program and 241 students were enrolled in the program in summer 2017. Scholars who attended 80% or more of the Summer Program were included in the treatment group (20 of 25 days offered in summer 2016 and 19 of 24 days offered in summer 2017). In summer 2016, 103 BellXcel scholars met the requirements for inclusion in the treatment group and 106 scholars met the requirements in summer 2017.

To build the body of evidence for the model, the BellXcel Summer Program participated in two random assignment evaluation studies in 2005 and 2012. In 2005, the Urban Institute evaluated the efficacy of the program for 835 elementary school students in New York and Boston (Chaplin & Capizzano, 2006). Students in the treatment group who participated in BellXcel gained about one month's worth of reading skills more than students in the comparison groups who did not attend BellXcel's programming. The study also found evidence of positive impacts on the degree to which parents of students in the treatment group encouraged their children to read. Further, in 2012, BellXcel was evaluated a second time by MDRC with grant funding from the Social Innovation Fund and the Edna McConnell Clark Foundation (Somers, Welbeck, Grossman, & Gooden, 2015). This second study measured the impact of the BellXcel Summer program on the reading and math skills of 919 rising 6th, 7th, and 8th graders in three school districts. The study found that BellXcel scholars had stronger math skills than students in the comparison group who did not attend the program—indeed about as large as might be expected from a five-week program during the regular school year. However, the results were not statistically significant and the BellXcel students did not demonstrate stronger reading skills. The lack of statistical

significance for math was due in part to the fact that the standard errors were significantly larger than anticipated.

The BellXcel Summer Program's incoming level of evidence was preliminary and this study targeted a moderate level of evidence. A multi-site randomized controlled trial of a similar summer program BellXcel implemented in prior years found a positive significant effect of participation on student academic achievement. Since the similar program was not identical to the program implemented through OnTrack Greenville, the incoming level of evidence was only preliminary instead of moderate. The Sub-Grantee's experience participating in rigorous evaluation and the program's promising results set a solid foundation for targeting a moderate level of evidence. With the availability of administrative data to measure student academic impacts in attendance, behavior, and course performance for students across the district and state, researchers were confident that a quasi-experimental design would provide more robust and technically sound results to expand the evidence base for the BellXcel Summer Program. However, due to the limited geographic scope of the initiative and the inability to randomly assign students to treatment and control conditions, researchers were not able to design a study to target a strong level of evidence.

In order to achieve a moderate level of evidence, this study utilized a single-site non-randomized group design with groups formed by propensity score matching. For confirmatory impact research questions, there were three comparison groups. Treatment students were matched to (1) other students in the *treatment schools* who did not participate in the intervention; (2) other students in the *same school district* attending four non-treatment non-Title I district schools; and (3) other students attending *Title I schools across the state* of South Carolina. The use of these multiple comparison groups improved the overall internal and external validity of the study, as each comparison group presented different threats to validity. Researchers matched students using a propensity score model that included race, gender, grade level, English proficiency, special education status, free and reduced meal eligibility, and baseline outcome variables. Researchers conducted separate matching procedures for each data source, administrative data and survey data. At the conclusion of the matching process, researchers ensured that there were no significant differences between the treatment and comparison groups on pre-treatment covariates. Table 1 below shows the final sample size numbers of all treatment and comparison groups.

To assess the impact of BellXcel on student attendance, as well as exploratory outcomes, researchers created seven matched comparison groups. It was necessary to create seven distinct, matched comparison groups due to (1) the three different comparison school populations (treatment schools, district schools, and state schools), (2) the two different sources of outcome data (administrative data and student survey data), and (3) two years of analysis (2016-17 and 2017-18).[1] Table 1 below shows the final sample size numbers of all treatment and comparison groups.

The study drew on administrative data and survey data to measure impacts and secondary outcomes. Through data-sharing agreements with Greenville County Schools and the South Carolina State Department of Education (SCDE), researchers received access to student administrative and test data to measure academic impacts. The primary intended impact of the BellXcel Summer Program course was to improve student course performance in math and English/language arts. Drawing on administrative data, researchers used the following measure of student course performance: *growth in MAP reading and math test scores.* Researchers administered a school-wide pre- and post-survey at treatment and within-

[1] Only 2016-17 administrative data were available for the state match.

district comparison schools to measure the BellXcel's exploratory secondary outcomes. This survey included seven validated measures of student perceptions of school climate and academic ability: *academic self-confidence, academic perseverance, valuing education, relationships with caring adults, relationships with teachers, school engagement, and school belonging.*

Table 1. Final Sample Size Numbers of Treatment and Comparison Groups

Year	Type of School Comparison Group	Group	Admin Data	Survey Data
Summer 2016	Treatment Schools	Treatment Students	100	100
		Comparison Students	226	226
	Within-District Schools	Treatment Students	100	100
		Comparison Students	226	226
	State Schools	Treatment Students	98	---
		Comparison Students	465	---
Summer 2017	Treatment Schools	Treatment Students	100	74
		Comparison Students	261	175
	Within-District Schools	Treatment Students	98	78
		Comparison Students	282	250

Note: This table shows the number of unique students matched for each comparison.

To answer confirmatory research questions, researchers compared the treatment and control groups on outcomes of interest to provide an estimate of the causal effect of participating in the BellXcel Summer Program. Researchers conducted multivariate regressions with the matched groups to allow for the inclusion of covariates to increase precision. Effect sizes and significance tests are presented for these results.

The ***confirmatory impact research question*** for the study was: *Did students who participated in the BellXcel Summer Program demonstrate improved math and English/language arts course performance when compared to matched comparison students, as measured by MAP assessment scores in reading and math?* Researchers were not able to confirm the hypothesis that BellXcel Summer Program students would have improved course performance when compared to matched comparison students. Key findings included:

- In summer 2016, there were no statistical differences in math or ELA course performance as measured by growth in MAP reading and math test scores when comparing BellXcel scholars to matched comparison students at treatment and district schools.
- MAP reading and math test scores were not available for treatment or comparison students in academic year 2017-18.

The ***exploratory impact research questions*** for the study were: *Did students who participated in the BellXcel Summer Program have fewer behavioral incidences than matched comparison students? Did students who participated in the BellXcel Summer Program have higher school attendance rates than matched comparison students?* Key findings included:

- Summer 2017 BellXcel scholars had higher average daily attendance in academic year 2017-18 than their matched counterparts attending treatment ($p < 0.01$) and district schools ($p < 0.01$).
- There were no statistically significant differences in chronic absenteeism between BellXcel scholars and matched comparison students during either academic year.

- Summer 2016 BellXcel scholars had fewer behavioral referrals than matched comparison students attending district schools during the academic year following program participation ($p < 0.10$).
- Across both years of the study, BellXcel scholars generally had significantly fewer days of out-of-school suspension than matched comparison students. Summer 2016 BellXcel scholars received 5.47 fewer days of out-of-school suspension than matched students at district schools ($p < 0.01$). Summer 2017 BellXcel scholars received 0.73 fewer days of out-of-school suspension than matched students at treatment schools ($p < 0.05$) and 0.59 fewer days than matched students at district schools ($p < 0.05$).
- BellXcel scholars also had fewer hours of in-school suspension than matched students at district schools. Summer 2016 BellXcel scholars had 0.72 fewer hours of in-school suspension ($p < 0.10$) and summer 2017 BellXcel scholars had 2.91 ($p < 0.01$) fewer hours of in-school suspension than their matched counterparts at district schools.

The ***exploratory research questions related to secondary outcomes*** were: *Did students who participated in the BellXcel Summer Program demonstrate higher self-confidence than matched comparison students? Did students who participated in the BellXcel Summer Program demonstrate a more positive attitude towards learning than matched comparison students? At the conclusion of the BellXcel Summer Program, did students who participated in the program demonstrate an increase in reading and math skills as measured by the STAR assessment scores? At the conclusion of the BellXcel Summer Program, did the parents of students who participated demonstrate an increase in engagement in their children's education?* Again, results varied by outcome and academic year. Key findings included:

- When compared to matched comparison students, BellXcel scholars did not have higher levels of self-confidence or an improved attitude toward learning after participating in the BellXcel Summer Program. Summer 2017 BellXcel scholars had significantly lower levels of self-confidence than matched comparison students at district schools, as measured by student survey scales of Academic Self-Confidence ($p < 0.05$) and Academic Perseverance ($p < 0.10$).
- Scholar growth in math and reading varied by program year, as measured by an internal program assessment. In summer 2016, BellXcel scholars did not show any growth in math or reading skills over the course of the program in summer 2016. In Summer 2017, however, BellXcel scholars did demonstrate significant growth in math ($p < 0.10$) and reading skills ($p < 0.01$).
- The majority of parents of BellXcel scholars reported on a post-program survey that they were more involved in their student's learning since enrolling their student in the BellXcel Summer Program. In summer 2016, 93% of parents reported that they were more involved in their child's learning. In summer 2017, 78% of parents reported that they were more involved.

The ***additional exploratory research questions related to secondary outcomes*** were: *Did students who participated in the BellXcel Summer Program report improved relationships with their teachers? Did students who participated in the BellXcel Summer Program report improved relationships with caring adults in their school? Did students who participated in the BellXcel Summer Program report improved school engagement? Did students who participated in the BellXcel Summer Program report participating in educational activities over the summer at a higher frequency than comparison students?*
Key findings included:

- When reflecting on their summer activities on a post-program survey, Summer 2016 BellXcel scholars more frequently reported playing math games over the summer than matched comparison students attending district schools ($p < 0.01$). Summer 2017 BellXcel scholars more frequently reported writing ($p < 0.01$), playing math games ($p < 0.05$), and going to a community center or camp ($p < 0.10$) over the summer than matched students at district schools. They also reported less frequently playing on a phone, watching TV, or playing video games than district school matches ($p < 0.05$).

- In general, there were no significant differences in school engagement or relationships with adults at school between BellXcel scholars and matched comparison students at either school group. The only exception was that summer 2017 BellXcel scholars did report lower levels of school engagement than matched students attending treatment schools ($p < 0.10$).

The ***implementation research questions related to program input and activities*** were: *Were program components (such as staffing, curriculum, data systems, provision of transportation, and community engagement events) implemented as intended at each school? Did the quality of the program reflect the intended design? What aspects of the design were modified at each school, if any, and why were they modified? What were the barriers to implementation, if any?* Key findings included:

- During the summers of 2016 and 2017, BellXcel implemented its Summer Program in OnTrack Greenville middle schools with a relatively high degree of fidelity to the BellXcel logic model.
- There was relatively high degree of alignment between the intended and actual implementation of BellXcel at each of the three OnTrack Greenville middle schools.
- Teachers modified the math curriculum to better meet the needs of the program scholars as some required more (or less) time to understand the concepts.
- There were no barriers to implementation.

The ***implementation research questions related to teacher activities and outputs*** were: *Did 100 percent of teachers, teaching assistants, and enrichment assistants receive the proper dosage of training? To what extent did the training provide teachers and assistants with the necessary skills and knowledge they needed to implement the program? What additional knowledge or skills did teachers and assistants require, if any? To what extent were teachers able to find support to fill in knowledge or skill gaps? To what extent did teachers implement the curriculum as intended? What challenges emerged as teachers implemented the curriculum? What suggestions do teachers have for improvement of the various program components as implemented?* Key findings included:

- 100% of teachers completed the in-person BellXcel training in summer 2016 and summer 2017. In summer 2016, the online training course was mandatory and 88% of staff completed at least 90% of this training. In summer 2017, the online training course was optional, with content covered during the in-person training, and 68% of staff completed at least 90% of the online course.
- Teachers implemented elements of the program, including the enrichment curriculum, community time, and field trips, with more alignment to the program design in 2017 than in 2016.
- Teacher survey responses indicated that they were satisfied with the training they received.
- Teachers found it challenging to teach the BellXcel curriculum, with its set structure and pacing, to scholars with a variety of needs.
- Suggestions included: having different training tracks for new and returning BellXcel teachers, simplifying the lesson plan template and process, making the pacing guides more realistic and less structured, acknowledging that teachers can use their professional judgment to make adjustments, and focusing more on helping teachers differentiate instruction for the wide range of scholar needs.

The ***implementation research questions related to scholar attributes, dosage, activities, and outputs*** were: *How were scholars recruited at each site and to what extent was the Early Warning and Response System (EWRS) used to identify scholars? What percentage of scholars who participated in BellXcel at each site met the original target characteristics specified in the model (at or below the 40th percentile in ELA or math? Were scholars provided the proper dosage of the program? Of those who completed the program, was their average daily attendance 80% or higher? What barriers, if any, prevented scholars from attending the program and/or using transportation at the desired rates? What aspects of the*

program did scholars prefer over others? What suggestions did scholars have for improvement of the various program components as implemented? Key findings ingluced:

- 5th grader students who were at or below the 60th percentile in ELA or math were recruited to attend the BellXcel Summer Program in summer 2016 and summer 2017.
- Across both summers and all school sites, the average daily attendance was 80%.
- Common reasons for missing days of the BellXcel program included: sickness, family vacation, and doctor's appointment or other appointment. Transportation to the program was a barrier for a small part of the scholar population.
- At least 90% of students enjoyed the field trips, and at least 85% enjoyed enrichment classes. Scholars preferred the math classes slightly more than the ELA classes.
- Scholars suggested having more "activities" and less "classroom time."

The ***implementation research questions related to parent engagement activities and outputs*** were: *Did at least 70% of parents attend parent engagement events? How effective were the parent engagement activities? How might parent engagement events be improved?* Results varied by site and year. Key findings included:

- No site reached 70% parent participation at any one event in either 2016 or 2017.
- At least 70% of parents from both years of implementation agreed that events helped them to understand their scholars' experience in the BellXcel program, were well-planned and had a clear agenda, and provided the opportunity for them to connect with BellXcel teachers and staff.
- Suggestions for improvement included having the engagement events on a different schedule so more parents can attend and to have a call or letter home to inform parents about the events.

Researchers were not able to confirm the hypothesis for the confirmatory research question. Summer 2016 BellXcel scholars did not have significantly different course performance in math and ELA than matched comparison students, as measured by the MAP assessment. When the school district stopped administering the MAP assessment in fall of 2017, researchers no longer had access to the confirmatory impact outcome measure for this study. As the evaluation continues for an additional two years, researchers, district staff, and project stakeholders will continue to look for alternative measures of math and ELA course performance that are available for both treatment and comparison students.

Exploratory impact results revealed positive significant results of the BellXcel Summer Program in student attendance and behavior, allowing the study to achieve a moderate level of evidence. While improving student attendance and behavior were not explicit impacts of the program identified in the program logic model, researchers conducted these additional analyses to expand the body of evidence for the program. For these outcomes, including average daily attendance, the number of hours of in-school suspension, and the number of days of out-of-school suspension, treatment students had significantly better outcomes than matched comparison students, primarily those students attending district schools. While it is possible that other factors influenced these findings, such as school-wide culture and policy change efforts related to OnTrack Greenville, they are important results and worth additional exploration in future years of the study.

There were several limitations to this study. Above all, the unavailability of MAP assessment scores for academic year 2017-18 prohibited researchers from conducting confirmatory impact analyses for summer 2017. While the use of multiple comparison groups generally was a strength of the study, it is worth emphasizing that the district comparison schools were not Title I schools. Though the district schools shared the same district and community context as treatment schools, the learning environment

at these schools likely was different; therefore, it is important to interpret results for the district school matched comparison students with caution. Results from the post-program student survey indicated that some matched comparison students may have participated in summer learning experiences that researchers were not able to account for in the matching procedure. While there were no other summer programs similar to the BellXcel Summer Program in the school district, any exposure to summer experiences could have influenced results. Lastly, the sample size of some analyses was lower than anticipated, as fewer students than expected attended at least 80% of the program.

As the study continues without SIF funding, researchers will consider additional measures for academic and social-emotional outcomes. Researchers also will examine the influence of participation in the BellXcel Summer Program over multiple summers, as the program recently dedicated a small number of spaces to rising seventh-grade students who participated in the program the prior year. Program staff will continue to focus on improving student attendance at the program, which will increase the sample size of the study and improve the likelihood of detecting significant results.

There were few key updates to the evaluation timeline, budget, program, or research team. The major update was the change in evaluation timeline due to the lack of Social Innovation Fund continuation funds to complete the final two years of program implementation and evaluation. As such, researchers executed a contingency plan to end the study after Year 3 (AY 2017-18). Members of the research teams at the Riley Institute at Furman University and RTI International remained constant, but there were some changes to BellXcel staff members managing the project. In January of 2017, BellXcel shifted project management from its regional Carolinas office to a dedicated Program Director for its OnTrack Greenville sites. There was some turnover among site leaders each summer, which was not unexpected.

This final report satisfies evaluation requirements for United Way of Greenville County's Social Innovation Fund grant award. Local leaders have committed to funding the initiative and evaluation for the final two years of the project in the absence of Social Innovation Fund continuation funding; therefore, evaluation next steps include the continuation of data collection and analysis as planned for academic years 2018-19 and 2019-20. Researchers will begin to disseminate preliminary results as early as 2019, but expect final results and a more robust dissemination plan to be available in March 2021.

I. Introduction

This report describes the implementation and impact evaluations of the BellXcel Summer Program, a Sub-Grantee intervention within United Way of Greenville County's SIF-funded OnTrack Greenville initiative. This is a final report submitted to the Social Innovation Fund to satisfy grant evaluation requirements and it addresses all implementation and impact research questions from the Sub-Grantee Evaluation Plan (SEP). The intended audience of this report is the Social Innovation Fund as well as Grantee and Sub-Grantee stakeholders.

Leaders from nonprofits, the school district, and the community implemented OnTrack Greenville, a collective impact approach that includes the implementation of an Early Warning and Response System (EWRS) in four target middle schools. The EWRS uses real-time data to identify and flag students at-risk of disengaging from school. An EWRS team, also known as an OnTrack Team, meets weekly and includes a team of educators and student support specialists who discuss the unique needs of identified students and match them with appropriate response interventions, tracking each student's progress over time. OnTrack Greenville's federally supported Social Innovation Fund (SIF) portfolio funded five Sub-Grantee interventions to ensure students have access to evidence-based interventions and supports. These five interventions include (1) a summer learning program for rising sixth-grade students; (2) integrated student support services; (3) a semester-long character development course; (4) school-based health centers; and (5) literacy coaching for teachers. This report examines one of these subgrantee interventions: the BellXcel Summer Program.

A. Program Background and Problem Definition

1. Description of Community and Program Need

Since United Way of Greenville County applied for this Social Innovation Fund grant in 2014, the local community has continued to experience significant growth and development. After the biennial census in 2010, the population of Greenville County grew by 12.7% to more than 500,000 people.[2] With a blossoming downtown, the city of Greenville has appeared on several national lists of best cities to live in or visit (Walker, 2018). The unemployment rate in the county dropped from 5.6% in February of 2014 to 2.5% in May of 2018.[3] At the same time, the county-wide poverty rate has decreased from 15.2% in 2014 to 12.4% in 2018.[4] A broad look at community indicators suggests many county residents are experiencing improved economic conditions.

A closer look reveals that not all residents have shared in this growth, especially in the White Horse Community, the geographic area targeted by OnTrack Greenville. As community developers have worked to revitalize neighborhoods close to the city center, low-income residents have continued to relocate to the White Horse Community, which straddles the edge of the city of Greenville. A recent assessment of neighborhood needs and assets revealed that many neighborhoods located in the White Horse Community, despite their wealth of community assets, continue to face challenges with

[2] U.S. Census Bureau 2018 Population Estimates
[3] U.S. Department of Labor 2018 Labor Force Statistics
[4] U.S. Census Bureau 2018 Small Area Income and Poverty Estimates (SAIPE)

unemployment, family poverty, income inequality, housing, and access to healthcare and childcare, among others (Cohen et al., 2017).

Public schools in the White Horse Community are part of Greenville County Schools. The largest district in the state of South Carolina and 45th largest district in the nation, Greenville County Schools consists of 101 schools and centers serving 76,900 students with 6,000 teachers. Approximately half of Greenville County Schools students are living in poverty (52%) and/or eligible for free or reduced-price meals (52%).

OnTrack Greenville serves four middle schools located in the White Horse Community. These middle schools serve a higher proportion of low-income and minority students than other schools in the district. In academic year 2017-18, each of these OnTrack Greenville sites had at least 79% of students living in poverty and 100% of students eligible for free or reduced-price meals. Three of the four treatment middle schools receive Title I funds, while the fourth site is technically a school program and ineligible for Title I funds despite a high proportion of students living in poverty. Across these three sites, the Title I funds have been used for expenses such as: teacher salaries, instructional technology, instructional materials, social workers, nurses, parent and family engagement coordinators, translators, tutoring, and other student services. The Title I funding also can enable schools to reduce the size of some classes and provide additional support staff.

The demographic characteristics of OnTrack Greenville treatment school student populations varied from the characteristics of the entire district[5]. In academic year 2017-18, OnTrack Greenville treatment schools were home to a high percentage of Hispanic or Latino students. The percentage of Hispanic students attending OnTrack Greenville treatment schools ranged from 27% to 55%, higher than the district average of 18%. In addition, OnTrack Greenville schools generally had a higher percentage of Black or African American students (23% to 55%) than the district average of 23%. OnTrack Greenville schools also had a higher poverty index than the overall district poverty index. The percentages of male and female students attending OnTrack Greenville treatment schools were reflective of the district average.

Table 2. School Enrollment by Gender, Race or Ethnicity AY 2017-18, 180th Day

Site	Enrollment (2017-18)	Gender		Race/Ethnicity				Poverty Index
		F	M	Black	White	Hispanic	Other	
District	75,220	49%	51%	23%	54%	18%	8%	53
Treatment School - Maximum	746	55%	57%	55%	26%	55%	9%	88
Treatment School - Minimum	109	43%	45%	23%	16%	27%	1%	79

One key academic indicator for predicting early disengagement among middle school students is course performance (Balfanz & Fox, 2011). Overall, students attending OnTrack Greenville middle schools placed well behind their peers on the South Carolina standardized assessment in ELA and math (SC READY) in academic year 2017-18. As shown below in Table 3, the percentage of students who met or

[5] Greenville County Schools Population Statistics 2017-18 180th Day Enrollment Summary https://www.greenville.k12.sc.us/About/main.asp?titleid=statistics1718

exceeded state standards in ELA ranged from 6.3% to 25.7% at OnTrack Greenville schools, while the percentage of students who met or exceeded state standards in math ranged from 3.6% to 25.7%.[6] These ranges of scores were well below the district and state averages in both subject areas.

Table 3. Percentage of Students who Met or Exceeded State Standards in ELA and Math: AY 2017-18

	Number of students	SC READY ELA	SC READY Math
State of South Carolina	340,478	41.7%	44.6%
District	34,220	48.9%	52.5%
Treatment School - Maximum	654	25.7%	25.7%
Treatment School - Minimum	112	6.3%	3.6%

Source: SC School Report Cards, 2019

The transitions from elementary to middle school and middle to high school are critical turning points in the social and academic development of an adolescent. If a student already is struggling in school, these transitions may cause irreparable emotional and academic setbacks. Research indicates that without intervention, children from low-income communities lose an average of one to two months grade-equivalent skills in math and reading each summer (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Entwisle & Alexander, 1992). This summer learning loss comprises up to two-thirds of the academic achievement gap between children from low-income backgrounds and their higher-income peers (Alexander, Entwisle, & Olson, 2007a). As years pass, cumulative summer learning loss contributes to chronic academic under-performance and risk of dropping out of school (Alexander, Entwisle, & Olson, 2007b).

In order to combat summer learning loss and improve academic performance among low-income students in the three target middle schools in the White Horse Community in the summers of 2016 and 2017, BellXcel operated a rigorous and stimulating six-week summer program for approximately 240 rising sixth grade students (80 at each school) each summer who were identified as academically and/or behaviorally at-risk. For four days each week, enrolled students received three hours of academic instruction, three hours of enrichment courses, free transportation, a healthy breakfast and lunch, and the opportunity to engage in recreational activities and go on field trips. In total BellXcel offered 25 days of the Summer Program in 2016 and 24 days in 2017. BellXcel's model for engaging with each school is designed to create a seamless connection between the summer program and the academic year. By the end of six weeks, these rising sixth graders are expected to have made gains in self-confidence and academic success, readying them to succeed in their core academic subjects during the following school year. A full description of the program model follows.

2. Description of Program Model

BellXcel, previously known as Building Educated Leaders for Life (BellXcel), was started in the 1990s by a group of Black and Latino students at Harvard Law School, led by Earl Martin Phalen and Andrew L. Carter, who saw a need for a comprehensive program which could help alter children's trajectories. Over the past 25 years, BellXcel has become a national leader in summer and afterschool programming. Alongside students and parents, BellXcel educators strive to help students excel through academic engagement, the development of social-emotional skills, and increasing self-confidence.

[6] South Carolina Department of Education 2018 South Carolina College- and Career-Ready Assessments (SC READY) Test Scores https://ed.sc.gov/data/test-scores/state-assessments/sc-ready/2018/

BellXcel's approach to increasing middle school student engagement is a six-week academic summer program to target "summer learning loss," the academic regression that occurs for many low-income students during summer vacation. The BellXcel Summer Program is a six-week, full-day learning experience that combines academic instruction with fun and hands-on enrichment activities, field trips, and service projects. The program is designed to help students gain new academic and social skills, improve student self-confidence and attitudes toward learning, and engage parents in the summer learning experience. In the fall following the program, the goal is for students to demonstrate improved math and ELA course performance.

The program achieves these impacts by hiring and training certified teachers and teaching assistants (Augustine, McCombs, Schwartz, and Zakaras 2013; McCombs, Sloan, Kirby, and Mariano 2009) to deliver small-group instruction in literacy and math (Cooper, Charlton, Valentine, Muhlenbruck, and Borman 2000; Augustine et al. 2013; McCombs et al. 2009). BellXcel's staff use research-based curricula aligned with Common Core standards (Borman, Benson, and Overman 2005; Funkhouser, Fiester, O'Brien, and Weiner 1995), and apply data from computer adaptive assessments to differentiate instruction according to students' unique learning needs (Cooper et al. 2000). In addition to academic instruction, students also experience highly engaging enrichment courses and activities focused on topics such as STEM (science, technology, engineering, and math), creative arts, and health and exercise (Birmingham, Pechman, Russell, and Mielke 2005). BellXcel works to involve parents in the program by providing feedback on program activities and their child's progress, and by engaging them in celebrations of their child's work (Cooper et al.2000; Funkhouser et al. 1995). And lastly, BellXcel provides free transportation and uses innovative incentives to encourage high student attendance rates (Borman et al. 2005; Borman and Dowling 2006).

The following section describes in detail the BellXcel program model as implemented through OnTrack Greenville, including the number of participants, inputs, activities, outputs, and key outcomes as outlined in the BellXcel logic model (Appendix B).

Program Inputs

As shown in the logic model, the implementation of BellXcel involved six different inputs: (1) BellXcel program staff, including certified teachers, teaching assistants, enrichment assistants, and instructional coaches; (2) schools and staff that collaborate with BellXcel and its programs; (3) a Common Core aligned curriculum, classroom space, and supplies; (4) financial resources from UWGC Social Innovation Fund subgrant, Greenville Partnership for Philanthropy, and other match sources; (5) Early Warning and Response System and internal data systems to identify eligible students, track students' progress, and provide ongoing feedback; and (6) OnTrack Greenville collective impact resources and support.

> (1) Staff members and positions critical to the successful implementation of BellXcel Summer programming included the BellXcel Executive Director for the Carolinas, the Director of Program Operations, the Director of Evaluation, three Program Managers, three Program Assistants, three Instructional Coaches, twelve Academic Teachers, twelve Enrichment Teachers, and twelve Teaching Assistants. Additionally, when working in schools with high English Language Learner (ELL) populations, such as School 1 and School 3, BellXcel recruits bilingual staff to help with parent and community engagement. BellXcel also provides ELL training for other staff members when possible.

(2) BellXcel prioritizes its connection with the administration and staff of the schools in which it serves. BellXcel hires certified teachers and staff from the local districts in order to increase cooperation, build local capacity, and ensure that the intervention is locally appropriate. In addition, teachers have the opportunity to provide feedback on the BellXcel experience through surveys at the end of the program. This feedback process is essential for collaboration and mutual respect between the BellXcel staff and the school administrators.

(3) BellXcel utilized the Renaissance Learning STAR Enterprise Assessment in literacy and math to identify student academic levels and needs when they enter the summer program. STAR results helped BellXcel adapt students' learning to local and/or national Common Core standards. The STAR results also were used to direct the development of curriculum, teacher training courses, and student individual instruction.

(4) The successful implementation of the BellXcel Summer program in each of the three target schools required $474,344 per fiscal year**.** The Social Innovation Fund grant contributed $232,961, with the remaining $241,383 coming from match sources in the Greenville Community.

(5) The EWRS allows schools to quickly and regularly identify students who are at risk of, or who already are, sliding off track. Using a color-coded dashboard, the EWRS helps school staff members easily recognize early warning signals for their students, such as a decrease in attendance or increase in disciplinary incidences. In addition, the data provided by the EWRS can be aggregated to show trends across the school, across grade levels, and across student subgroups. By giving teachers and other school staff members access to real-time data related to attendance, behavior, and course performance, the EWRS is an essential tool for assessing individual students' strengths and needs, and then using that information to provide students with the appropriate interventions.

(6) OnTrack Greenville is a community-wide initiative to ensure middle school students stay on track toward high school graduation. The initiative works with treatment schools, implementation partners, nonprofits, community members, government officials, funders, and other stakeholders to achieve the common goal of keeping students on track towards high school graduation and future success. OnTrack Greenville consistently convenes school leadership, implementation partners, and funders to coordinate and implement key aspects of the initiative for the coming school year, while also building a shared vision, governance, and accountability for OnTrack Greenville. Engaging with the community, families, students, other nonprofits, and grassroots organizations contributes to the overall collective impact of the initiative.

Activities and Outputs

With these six inputs, BellXcel Summer Program conducts the following activities: (1) training for teaching and enrichment assistants; (2) offering a six-week summer program operating 6.5 hours per day for 4 days each week; (3) increasing community and parent engagement through engagement events; (4) giving daily instruction consisting of 3 hours of academic instruction and 2 hours of enrichment courses; and (5) providing daily free transportation to and from the summer program.

(1) All leadership and teaching staff (including enrichment teachers) completed a combination of online and in-person training. The online training, which was mandatory in 2016 and optional in 2017, consisted of up to six hours of interactive on-line training through BellXcel's unique "BellXcel University" e-learning platform. All leadership and teaching staff received between 16-24 hours of in-person training prior to the start of the program. The in-person leadership and staff training provided more in-depth training on the BellXcel program model and effective site operations, teaching practices, data-driven instruction, and family engagement.

(2) The six-week summer program provided daily instruction for 6.5 hours per day, four days per week. **Each day of the summer program is the program unit.**

(3) At BellXcel, parents are valued as equal partners in their student's learning. Information on student progress was given formally and informally to parents in the forms of progress reports, phone calls, and conferences. Each BellXcel Summer Program site was required to have at least one family event where parents were invited to the school site, in addition to a culminating Closing Ceremony event celebrating and reflecting on the students' summer learning experience. BellXcel encouraged their sites to be extremely responsive to parents' questions and concerns, and information from parent surveys were used to help BellXcel refine the program.

(4) Daily instruction consisted of three hours of academic instruction and two hours of enrichment courses. Academic instruction time was shared between math and ELA instruction in the Common Core aligned curriculum. Enrichment courses varied by site and were aligned with students' interests. Enrichment course topics included coding, STEM courses, nutrition, dance, social skills, technology, cooking, art, and karate.

(5) In partnership with Greenville County Schools, BellXcel offered free bus transportation to and from the summer program every day. Parents and scholars who requested bus transportation were added to the transportation bus route for each school site. The provision of free transportation to and from the program helped remove potential barriers to attending the program.

With the activities described above, the implementation of the BellXcel Summer Program was expected to result in the following outputs: (1) 100% of certified teachers, Teaching Assistants, and Enrichment Assistants trained; (2) students who remain enrolled and complete the program have 80% or higher average daily attendance in the program; (3) 80% or more of enrolled students utilize free transportation to and from the summer program; and (4) 70% of parents attend parent engagement events.

Outcomes and Impacts

The exploratory outcomes which BellXcel sought to achieve included: (1) increasing student participants' reading and math skills; (2) increasing their self-confidence; (3) improving their attitude toward learning; and (4) increasing their parents' engagement in their education. As a result, the primary confirmatory impact of BellXcel was improved math and ELA course performance at the beginning of the academic year through the prevention of summer learning loss. The exploratory impacts of BellXcel included increased student attendance and improved student behavior. Each of the outcomes affects the confirmatory and exploratory impacts in the following ways:

(1) The three hours of academic instruction during the six-week summer program was expected to lead to an increase in student reading and math skills by the end of the program.

(2) BellXcel's focus on improving student reading and math skills and giving students the opportunity to gain new social skills through enrichment activities, field trips, and service projects was expected to lead to an increase in self-reported self-confidence. Parent engagement in and celebration of student achievements during the program also should positively affect student self-confidence.

(3) By working with students three hours each day over the course of the six-week program and by exposing students to fun and engaging enrichment activities and field trips, BellXcel expected to see an improvement in students' attitudes towards learning.

(4) By providing continuous opportunities for parents to view their children's work over the summer, provide feedback on activities and progress, and celebrate their children's accomplishments, BellXcel hoped to increase parents' engagement in their children's education.

This report examines all confirmatory and exploratory impacts as well as exploratory secondary outcomes.

B. Overview of Prior Research

The BellXcel Summer Program's theory of change is supported by a range of research and best practice reports on after-school programs. Several components of the BellXcel Summer Program are aimed at promoting student attendance, with a goal of achieving an average daily attendance rate of 80% or higher at each site. Prior research has helped confirm the importance of high attendance rates at summer programs. For example, two studies that examined the effects of the Teach Baltimore Summer Academy program on preventing summer learning loss on elementary school children found that regular attendance during the program was necessary in order to achieve improved academic outcomes (Borman, Benson, & Overman, 2005; Borman & Dowling, 2006). In addition, a study of voluntary summer learning programs found that students on average attended approximately 75% of the days offered by a program and that high levels of program attendance, defined as attending 20 days or more of a program, was related to short-term benefits in math and short-term and long-term benefits in ELA (Augustine et al., 2016).

The BellXcel model places a strong focus on training program administrators, academic teachers, enrichment teachers, and other program staff, a practice supported by prior research. McCombs et al. (2009) used a mixed-methods longitudinal study design (including interviews and surveys of principals) to look at two New York City (NYC) programs (Saturday Preparatory Academies and Summer Success Academies) that were part of NYC's 5th-grade promotion policy in 2007. Principals who participated in the study reported that lack of teacher program knowledge and experience hindered the success of these programs at some of the schools in the study. In addition, Sanders and Horn (1998) studied the impact of teacher effectiveness on student academic outcomes by utilizing school-housed, student-level data and teacher effectiveness data gathered utilizing the Tennessee Value-Added Assessment System methodology. They found that more than race, socioeconomic status of parents, or class size, teacher effectiveness was the largest determinant of student academic growth. Finally, Augustine et al. (2013) drew on over 1,800 surveys, 325 interviews, and 400 hours of classroom and enrichment activity

observations to evaluate summer programs in six school districts across the country in the summer of 2011. In the report based on their findings, Augustine et al. wrote "[h]iring effective teachers and giving them the [training] support they need are critical steps to maximizing student achievement" (p. xiv).

The importance of including enrichment activities in a summer program curriculum also has a strong basis in prior research. For example, Augustine, et al. (2013) and Birmingham, et al. (2005) reported that providing students with enrichment activities outside of the classroom was a consistent part of successful after-school projects. Specifically, Birmingham, et al. (2005) identified ten after-school projects that were high-performing and funded by The After-School Corporation (TASC) after analyzing student performance data over time from areas in which the TASC was being implemented. They found that while student academic performance was used to identify these ten after-school projects, the projects themselves did not focus primarily on academic growth. Rather, these ten projects focused on providing enrichment activities that acted as the first exposure to new worlds of art, dance, and organized sports for many of their participants.

C. Overview of Evidence Level and Impact Study

The BellXcel Summer Program entered the OnTrack Greenville SIF portfolio with a preliminary level of evidence. The BellXcel Summer Program participated in two random assignment evaluation studies in 2005 and 2012. In 2005, the Urban Institute evaluated the efficacy of the program for 835 elementary school students in New York and Boston (Chaplin & Capizzano, 2006). Students in the treatment group who participated in BellXcel gained about one month's worth of reading skills more than students in the comparison groups who did not attend BellXcel's programming. The study also found evidence of positive impacts on the degree to which parents of students in the treatment group encouraged their children to read. Since this program model and age of participants varied from program implementation through OnTrack Greenville, these results constituted a preliminary level of incoming evidence. Further, in 2012, BellXcel was evaluated a second time by MDRC with grant funding from the Social Innovation Fund and the Edna McConnell Clark Foundation (Somers et al., 2015). This second study measured the impact of the BellXcel Summer program on the reading and math skills of 919 rising 6th, 7th, and 8th graders in three school districts. The study found that BellXcel scholars had stronger math skills than students in the comparison group who did not attend the program—indeed about as large as might be expected from a five-week program during the regular school year. However, the results were not statistically significant and the BellXcel students did not have stronger reading skills. The lack of statistical significance for math was due in part to the fact that the standard errors were significantly larger than anticipated.

This study targeted a moderate level of evidence by utilizing a single-site non-randomized group design with groups formed by propensity score matching. The impact evaluation aimed to build additional evidence for the program's impact on student course performance and examine other exploratory impact variables, such as student attendance, student behavior, and other social-emotional outcomes related to character development and school success.

Researchers were not able to target a strong level of evidence for several reasons. First, the geographic scope of OnTrack Greenville was not large enough to support a national- or state-wide multi-site research design that typically is required to achieve a strong level of evidence. Second, the Early Warning and Response System (EWRS) used to identify and match students to interventions did not lend

itself to conditions in which randomization was feasible. The EWRS uses a wrap-around approach in which educators match a student to an intervention given the student's unique early warning indicators and needs. The identification and matching process is time intensive and intended to provide the best array of services to students. Initiative stakeholders expressed ethical concerns about withholding treatment to identified students in order to support random assignment, as the collective portfolio-level goals of the initiative were to improve academic achievement and engagement for all identified students at the target schools. Further, the number of students potentially identified for treatment through the EWRS was relatively small. Randomly assigning identified students to treatment and control groups would have decreased the sample size and threatened the study's statistical power.

For confirmatory impact research questions, there were three comparison groups. Treatment students were matched to (1) other students in the *treatment schools* who did not participate in the intervention; (2) other students in the *same school district* attending four non-treatment district schools; and (3) other students attending *Title I schools across the state* of South Carolina. The use of these multiple comparison groups improved the internal and external validity of the study, as each comparison group presented different threats to validity. Researchers matched students using a propensity score model that included race, gender, grade level, English proficiency, special education status, free and reduced meal eligibility, and baseline outcome variables. Researchers conducted separate matching procedures for each data source, administrative data and survey data. At the conclusion of the matching process, researchers ensured that there were no significant differences between the treatment and comparison groups on pre-treatment covariates.

Researchers assessed exploratory secondary outcome research questions using a school-wide pre/post-student survey at treatment and district comparison schools, with data collection occurring in October and May of each academic year. Since researchers were not able to administer the student survey to fifth-grade students in the spring before the Summer Program, October survey data serves as a post-only measure for the BellXcel Summer Program. There were approximately nine to ten weeks between the end of the Summer Program and the administration of the student survey. Six to seven of these weeks occurred during the academic year. The school-wide survey included the following exploratory secondary outcome measures: *academic perseverance, academic self-confidence, valuing education, relationships with teachers, relationships with caring adults, school engagement, and school belonging*.

D. Research Questions

The interim impact study includes the following types of research questions: (1) confirmatory impact research questions; (2) exploratory impact research questions; and (3) exploratory research questions related to secondary outcomes.

1. Impact Research Questions

a. Confirmatory Impact Research Questions

The question below is the *confirmatory impact research question*:

RQ1. *Did students who participated in the BellXcel Summer Program demonstrate improved math and English/language arts course performance, as measured by the MAP assessment, when compared to matched comparison students?*

b. Exploratory Impact Research Questions

The questions below are *exploratory research questions* related to the impacts:

RQ2. *Did students who participated in the BellXcel Summer Program have fewer behavioral incidences than matched comparison students?*

RQ3. *Did students who participated in the BellXcel Summer Program have higher school attendance rates than matched comparison students?*

c. Exploratory Outcome Research Questions

The next set of questions are *exploratory research questions* related to the *secondary outcomes*:

RQ4. *Did students who participated in the BellXcel Summer Program demonstrate higher self-confidence than matched comparison students?*

RQ5. *Did students who participated in the BellXcel Summer Program demonstrate a more positive attitude towards learning than matched comparison students?*

RQ6. *At the conclusion of the BellXcel Summer Program, did students who participated in the program demonstrate an increase in reading and math skills as measured by the STAR assessment scores?*

RQ7. *At the conclusion of the BellXcel Summer Program, did the parents of students who participated demonstrate an increase in engagement in their children's education?*

d. Additional Exploratory Research Questions

RQ8. *Did students who participated in the BellXcel Summer Program report improved relationships with their teachers?*

RQ9. *Did students who participated in the BellXcel Summer Program report improved relationships with caring adults in their school?*

RQ10. *Did students who participated in the BellXcel Summer Program report improved school engagement?*

RQ11. *Did students who participated in the BellXcel Summer Program report participating in educational activities over the summer at a higher frequency than comparison students?*

2. Implementation Research Questions

The implementation evaluation answered the following core research questions:

a. Focus Area #1: Program Inputs and Activities

RQ12. *Were program components (such as staffing, curriculum, data systems, provision of transportation, and community engagement events) implemented as intended at each school?*

RQ13. *Did the quality of the program reflect the intended design?*

RQ14. *What aspects of the design were modified at each school, if any, and why were they modified? What were the barriers to implementation, if any?*

b. Focus Area #2: Teacher Activities and Outputs

RQ15. *Did 100 percent of teachers, teaching assistants, and enrichment assistants receive the proper dosage of training?*

RQ16. *To what extent did the training provide teachers and assistants with the necessary skills and knowledge they needed to implement the program? What additional knowledge or skills did teachers and assistants require, if any? To what extent were teachers able to find support to fill in knowledge or skill gaps?*

RQ17. *To what extent did teachers implement the curriculum as intended? What challenges emerged as teachers implemented the curriculum?*

RQ18. *What suggestions do teachers have for improvement of the various program components as implemented?*

c. Focus Area #3: Scholar Attributes, Dosage, Activities, and Outputs

RQ19. *How were scholars recruited at each site and to what extent was the EWRS used to identify scholars? What percentage of scholars who participated in BellXcel at each site met the original target characteristics specified in the model (at or below the 60th percentile in English Language Arts or mathematics)? Which, if any, additional target characteristics were added during the recruitment phase? If additional criteria were added, what percentage of scholars met each of the target characteristics? What recommendation do district and BellXcel staff have for improving recruitment next year?*

RQ20. *Were scholars provided the proper dosage of the program (i.e., six weeks operating six and a half hours per day, four days per week, with three hours of academic instruction and two hours of enrichment)?*

RQ21. *What percentage of scholars completed the program at each program site? Of those who completed the program, was their average daily attendance 80 percent or higher?*

RQ22. *What barriers, if any, prevented scholars from attending the program and/or using transportation at the desired rates? What strategies encouraged attendance and transportation utilization?*

RQ23. *What aspects of the program did scholars prefer over others?*

RQ24. *What suggestions do scholars have for improvement of the various program components as implemented?*

d. Focus Area #4: Parent Engagement Activities and Outputs

RQ25. *Did at least 70 percent of parents attend parent engagement events?*

RQ26. *How effective were the parent engagement activities?*

RQ27. *How might parent engagement events be improved?*

E. Contribution of the Study

1. Level of Evidence Generated by the Study

This evaluation targeted a moderate level of evidence and examined the effects of completing the BellXcel Summer Program on multiple confirmatory impact measures of course performance and exploratory impact measures of attendance and behavior. Researchers adjusted for multiple comparisons in order to make a final determination of the evidence level this study achieved. As described below, statistically significant results for exploratory analyses on student attendance and behavior remained significant after adjusting for multiple comparisons ($p < 0.10$). Given the use of multiple comparison groups to reduce threats to internal and external validity and a successful matching process that yielded balanced treatment and comparison groups, this study's methodological rigor and positive significant results merit a moderate level of evidence for the BellXcel Summer Program.

For the confirmatory impact analyses examining course performance, there were four total tests (two each for the treatment and district comparison groups in academic year 2016-17). None of the tests found statistically significant differences between BellXcel and comparison students on course performance. However, scholars' math and reading skills, as measured by an internal pre- and post-program assessment, improved significantly over the course of the program in summer 2017. It is possible that this growth in skills would have translated into significant growth on the MAP assessment when compared to matched comparison students. The absence of MAP assessment data for treatment and comparison students in academic year 2017-18 was a serious challenge in this study. Researchers are working to identify an appropriate replacement outcome measure for future years of the study. While this is a final report to the Social Innovation Fund, these are interim impact findings, as program implementation and impact evaluation will continue through non-federal funding sources. Subsequently, it is possible that researchers will detect positive significant effects of program participation on course performance at the conclusion of the study.

For the exploratory impact analyses examining student attendance during the academic year following program participation, there were ten total tests (average daily attendance rate and chronic absenteeism for the five matches using the administrative data). Before adjustments for multiple comparisons were made, two of the tests indicated that BellXcel scholars demonstrated statistically significantly higher average daily attendance than comparison students. After adjusting for multiple comparisons, both results remained statistically significant. BellXcel scholars who attended the Summer Program in 2017 had higher rates of average daily attendance than matched comparison students at treatment and district schools in academic year 2017-18.

For the exploratory impact analyses examining student behavior during the academic year following program participation, there were 30 total tests (any referral, any in-school suspension, any out-of-school suspension, number of referrals, number of hours of in-school suspension, and number of days of out-of-school suspension for the five matches using the administrative data). Before adjustments for multiple comparisons were made, nine of the tests indicated statistically significant differences between BellXcel scholars and comparison students in the predicted direction. After adjusting for multiple comparisons, one result remained statistically significant. BellXcel scholars who completed the program in summer 2016 had fewer hours of in-school suspension than matched comparison students attending district schools in academic year 2016-17.

2. Strengths and Limitations of the Study

There were many strengths to this study. The use of multiple comparison groups improved the overall internal and external validity of the study, as each comparison group presented different threats to validity. The majority of positive significant findings were detected with the matched students attending district comparison schools. One strength is that these schools shared the same district and community context. Moreover, students in this comparison group were likely to have participated in BellXcel Summer Program if it had been available to them at their school. These schools did not share the same school or neighborhood contexts, though, presenting a threat to internal validity. As such, results for the district comparison group should be interpreted with caution.

In addition, a thorough implementation study strengthened the implementation of the program and allowed researchers to confirm a sufficient degree of model fidelity. The lessons learned through the implementation study were valuable to project stakeholders and helped shine a light on program strengths and possible areas of improvement.

However, there were several limitations to the study. The primary limitation was the unavailability of academic year 2017-18 MAP assessment scores for confirmatory impact analyses. Without this crucial outcome measure, researchers only were able to conduct confirmatory impact analyses for the program for summer 2016. Results from the implementation study indicated that the program was implemented with a higher degree of fidelity in summer 2017, while the internal pre-/post-program assessment of student math and reading skills showed significant growth in both subject areas in summer 2017. With the discontinuation of MAP assessment administration and the lack of a comparable pre-/post-program outcome measure for course performance, researchers were not able to assess if these improvements in model fidelity and math and reading skills translated into improved course performance.

In addition, researchers were not able to identify a subset of state Title I middle schools with student population demographics similar to the treatment schools. The Sub-Grantee Evaluation Plan called for

including only state comparison schools with a poverty index of 85% or higher and Hispanic students representing 10% of the student body. Only 13 schools met these inclusion criteria to be considered as state comparison schools—many were charter schools or special designation schools serving exceptional learners and were substantially different from the treatment schools. Researchers opted to loosen the inclusion criteria and include all Title I middle schools in South Carolina in the state school comparison group.

Another limitation of the study is that researchers did not have the ability to assess if comparison students at district and state schools participated in similar program services, such as other summer learning programs. While researchers could confirm that state comparison schools did not offer the exact BellXcel Summer Program, it is possible that a number of schools in the state comparison group offered similar summer programs for rising sixth-grade students. However, the number of matched students who attended these schools likely was very small and the inclusion of these students as matches would not have influenced the results of the study significantly.

In addition, the treatment schools were simultaneously implementing formal and informal school-wide initiatives to improve student academic outcomes. These school-wide efforts were confounding factors that may explain the lack of significant effects when comparing BellXcel students to in-school matched comparison students. These school-wide efforts also increased the likelihood that the positive significant effects of the program identified when examining district school matches may not be fully attributable to the BellXcel Summer Program.

Further, the absence of positive significant findings for in-school matches may be related to missing data on student participation in other OnTrack Greenville interventions. Apart from the school-wide models discussed above, OnTrack Greenville includes four other formal implementation partners and several informal partners, some of whom are working to improve the same student outcomes as the BellXcel Summer Program. It is possible that some of the in-school matches selected for the present study participated in other OnTrack Greenville support programs that influenced student outcomes. This study originally intended to control for participation in other OnTrack Greenville support programs to address this limitation; therefore, this represents a deviation from the Sub-Grantee Evaluation Plan.

In the evaluation presented here, the smallest number of BellXcel students included in the main confirmatory impact analyses was 100, slightly below the predicted sample size value included in the SEP power analysis. The small sample sizes in the evaluation limit the study's ability to identify a statistically significant effect of participating in the BellXcel Summer Program. With a larger sample size, some of the small effects detected by this study may have been statistically significant if there had been a larger sample. More information on the power analysis, minimum detectable effect size (MDES), and sample size appear later in this report.

3. Connection of this Study to Future Research

The results of this study highlight the challenges around measuring academic impacts in the field of summer learning. Studies that rely on secondary data sources for outcome measures, such as the present study, face limitations depending on the types of assessments school districts use and the timing of these assessments. While end-of-year standardized test scores are most readily available to researchers, these assessments occur after summer program participants have received an entire school year of academic instruction and are not an appropriate outcome measure for evaluating summer

learning programs. In future years of this study, researchers and program leaders will consider adjusting the outcome measures in order to assess more accurately the academic impacts of the program. Possible assessment tools include the Woodcock Johnson IV and a subscale of the MAP assessment. However, it would be challenging to administer these assessments with a sample of students attending treatment and district schools in order to have a pool of possible matched comparison students, requiring researchers to utilize a less rigorous impact study design. The field of summer learning research must continue to weigh the pros and cons of different measurement techniques and study designs as it engages in more dialogue around the expected academic outcomes of programs.

In addition to considering alternate measures of academic outcomes, future research should explore other measures of social-emotional outcomes. The measurement of social-emotional learning has been at the forefront of conversations in the fields of summer learning and out-of-school time and there are numerous assessment tools and toolkits available (e.g. American Institutes of Research, 2015). For programs like the BellXcel Summer Program that occur at middle schools and employ academic year teachers, improved relationships with teachers and academic self-confidence are important secondary outcomes leading to school success. Emphasis on these outcomes is essential for internal program performance management and external evaluations. In summer 2018, BellXcel piloted a new internal post-program student survey at OnTrack Greenville sites, the Holistic Student Assessment (HSA) from the PEAR Institute. As BellXcel assesses data collected from this pilot administration, there are future research opportunities to explore how scholars' responses to items on the HSA are correlated to their responses on the OnTrack Greenville Student Survey administered in the fall. These research efforts will contribute to the understanding of social-emotional outcome measurement in the field of summer learning.

There also are opportunities to examine the impact of participating in the program over multiple summers. In summer 2018, BellXcel enrolled a small cohort of rising seventh-grade students, all of whom attended the program the prior year as rising sixth-grade students. While no analyses of program data have been conducted, program staff perceived that these rising seventh-grader students attended the program with high levels of attendance and engaged deeply in the academic and enrichment activities. Researchers should examine the effects of participating in the program over multiple summers in order to inform future decisions around recruitment efforts and the structure of the program at OnTrack Greenville sites.

4. Changes to Sub-Grantee Evaluation Plan

The primary change to the SEP was the loss of a key outcome measure of course performance, MAP assessment scores in ELA and math. Prior to academic year 2017-18, the local school district administered the MAP assessment in grades 3 through 8 at least two times per year, in the fall and spring. Some schools opted to administer the assessment a third time, in winter. The district opted to end its contract with MAP and began administering Mastery Connect in fall of 2017. At the present time, researchers do not have access to Mastery Connect data and remain uncertain if data from this assessment will serve as an acceptable outcome measure in the study.

One modification was made to the treatment definition for the study. Previously, researchers proposed that students would need to be enrolled at least 51% of the academic year at the treatment schools in order to be included in the treatment group, excluding students who transferred to and attended a non-treatment school for the majority of the academic year. It proved challenging to measure this given the

structure of the attendance data; therefore, researchers removed this inclusion criterion from the treatment definition. The SEP also stated that "completion" of the BellXcel program would be part of the treatment definition. For this evaluation, researchers defined completion as attending at least 80% of the summer program. Few students attended every day of the program. The evaluation team considered 80% participation would be enough to produce possible impacts without being too high to limit sample sizes.

The outcome measures for the behavior variables were modified slightly. Originally in the SEP, researchers proposed including continuous behavior variables: number of behavioral referrals, number of hours of in-school suspension, and number of days of out-of-school suspension. In addition to these continuous measures of student behavior, researchers also added dichotomous categorical variables: any behavioral referral, any in-school suspension, and any out-of-school suspension.

Researchers modified the student survey outcome measure for the exploratory secondary outcome of student attitude toward learning. Researchers originally proposed using a four-item scale *Valuing School* that measured a student's beliefs about the importance of school (Rockman et al, 2013). Researchers also included a similar scale on the student survey from a prior evaluation of Communities In Schools (Corrin, Parise, Cerna, Haider, & Somers, 2015). After the first wave of data collection, researchers assessed the psychometric properties of both scales and ultimately opted to retain the entire latter scale with the addition of one item from the former scale. Researchers conducted exploratory factor analysis to examine the factor structure of the new scale, discussed later in this report in Section II.B.2.

There was a change in the timeline for receiving administrative data from the South Carolina Department of Education for the state comparison group analyses. Researchers anticipated receiving the state dataset in October or November for the prior academic year. However, the dataset for academic year 2016-17 was not available until February 2018 and researchers still do not have access to data from academic year 2017-18. This delay in receiving state data prohibited researchers from conducting the state comparison group analyses for this report.

Researchers also had to alter the inclusion criteria for state comparison schools. In the SEP, researchers originally proposed to select state schools with a poverty index of 85 or higher and a Hispanic student population of at least 10%. At the time of writing the SEP, South Carolina calculated the poverty index based on the number of students eligible for free or reduced meals. After the introduction of the community provision for free and reduced meals, state officials introduced a new measure of poverty that included students who met any of the following criteria: homeless or migrant during the academic year; Medicaid enrollment at any time during a three year period; SNAP enrollment at any time during a three year period; TANF enrollment at any time during a three year period; or foster care enrollment at any time during a three year period. This change in the poverty index affected and, in general, reduced the reported poverty levels of treatment schools and all schools across the state. When researchers searched for state comparison schools using these criteria, only 13 schools appeared as possible comparison schools from which to draw matched comparison students. Of these schools, several were charter schools or schools serving exceptional learners and were not appropriate to serve as comparison schools. As such, researchers relaxed the inclusion criteria for state comparison schools and included all Title I middle schools in the state of South Carolina outside of the local district.

In addition, researchers were not certain if they would be able to administer the OnTrack Greenville Student Survey outside of the treatment schools when preparing the SEP. The local district allowed researchers to administer this pre- and post-survey at the four district comparison schools, allowing

researchers to analyze student survey outcomes using matched comparison students at district schools. This change to the SEP strengthened the study's design for the analysis of secondary research questions. Survey administration proceeded as described in the SEP, though it was not possible to administer the pre-survey in September of each academic year due to the testing schedule and the amount of time needed to distribute parent opt-out letters before preparing survey materials. Instead, the student survey administration window occurred typically during the first two weeks of October each academic year of the study.

To examine possible attrition from the study, researchers proposed in the SEP assessing how students who attrited from the study through leaving the state dataset differed from students who remain in the study. Students were to be compared based on demographics, pre-treatment outcomes measures, and post-treatment outcome measures when possible. Evaluators proposed examining the possibility of differential attrition between treatment and control groups based on these factors as well. Researchers did not conduct these comparisons as proposed, as the structure of the data files and the matching procedure meant there were very few students for whom attrition occurred after students were matched. A comparison of the number of students matched and the number of students in each regression with the matched sample demonstrated that attrition of this type was not a widespread challenge for this study.

There were some additional modifications to the matching procedure. The SEP noted that researchers would trim observations with propensities less than 0.1 and greater than 0.9, if sample size permitted. This was not done to ensure larger sample sizes. In order to increase balance and overlap, researchers used matching with replacement, rather than matching without replacement which was specified in the SEP. This necessitated the use of frequency weights in the matched analyses.

Further changes were made to the impact analysis plan. While the SEP stated that the main analyses would focus on the "treatment-on-the-treated" (TOT) effect, the researchers also suggested an "intent-to-treat" (ITT) analysis might also be performed if data were available. The evaluation team did not have access to the EWRS data in a form that would allow the researchers to compare those who were identified to participate in the BellXcel program, as opposed to those who actually participated. Therefore, an ITT analysis was not possible. In addition, there were some challenges in estimating the proper standard errors in the regression analyses post-match. Researchers used a bootstrapping method, rather than clustered standard errors. The bootstrap standard errors were very similar to robust standard errors. Other modifications were implemented to maximize the number of observations to be included in the analyses.

The proposed final regression model in the SEP included pre-treatment, or baseline, measures of the outcome as a covariate. The analyses presented here used that approach for the attendance, behavior, and test score analyses. Researchers did not control for a pretreatment measure of the outcome for the survey analyses, as baseline survey data were not available.

In the SEP, researchers stated that they would use one-tailed tests with a significance level of $\alpha \leq .05$ to determine statistical significance. In this evaluation, researchers also considered the possibility that participating in the BellXcel Summer Program could decrease student academic performance. Therefore, researchers used a two-tailed, 0.10 alpha level. In terms of identifying a positive BellXcel effect to support a moderate level of evidence, the two approaches are equivalent.

This evaluation was complicated by the presence of multiple, simultaneous interventions occurring in the OnTrack Greenville schools. In the SEP, researchers suggested that participation in the other student-level interventions could be controlled for in the final regression models. However, involvement in these other programs could be influenced by BellXcel participation. Given that controlling for post-treatment covariates can bias estimates of causal impacts (Montgomery, Nyhan, & Torres, 2018), researchers did not adjust for participation in the other OnTrack Greenville programs. A full analysis of the effects of participating in different intervention combinations is better suited for the cumulative impact study, rather than this evaluation.

II. Study Approach and Methods

A. Impact Evaluation Design

This study used quasi-experimental methods to examine the effect of participation in the BellXcel Summer Program. While a randomized control trial (RCT) would have been ideal, it was not feasible for this study. Instead, researchers used nearest neighbor propensity score matching to estimate the effect of the BellXcel Summer Program on students at the three treatment schools. Matching techniques are popular in observational education research when an RCT is not feasible, and previous research has found that the results from matching can replicate RCT results if a number of assumptions are met (Bifulco, 2012; Cook, Shadish, & Wong, 2008; Fortson, Verbitsky-Savitz, Kopa, & Gleason, 2012).

This analysis compares BellXcel scholars to three different comparison groups. First, BellXcel scholars were matched to other students attending OnTrack Greenville middle schools who did not participate in the BellXcel program. This is the "treatment school" comparison group. A second matching procedure compared the BellXcel scholars to students in four other Greenville County Schools middle schools that were not a part of the OnTrack Greenville initiative. This is the "district school" comparison group. Finally, a third matching procedure compared the BellXcel scholars to other public school students attending Title I middle schools across the state of South Carolina. This is the "state school" comparison group. Table 4 presents a summary of these groups and the threats to internal validity posed by each group.

At the first stage of the matching process, the "treatment" was defined. Researchers used internal program records from the BellXcel Summer Program to generate a total pool of students who received services. Then, researchers limited inclusion in the treatment group to students in the treatment schools who met the following parameters: (1) the student's parent(s)/guardian(s) agreed to their child participating in the study, and (2) the student attended 80% or more of the days offered in the BellXcel Summer Program. In academic year 2016-17, researchers creating matches using a dataset with merged administrative and survey data, while in academic year 2017-18 researchers created separate matches for administrative and survey data. In academic year 2016-17, there were three comparison groups (non-BellXcel scholars in treatment schools, district schools, and state schools). In total, researchers produced seven sets of matches for this study.

Next, in order to estimate the true effect of the BellXcel Summer Program, researchers considered the observed and unobserved factors that may have affected both participation in the program and the outcomes of interest. This "first-stage" regression model included race, gender, free and reduced meal status, special education status, English proficiency, average daily attendance, total number of in-school suspensions, total number of out-of-school suspensions, a variety of standardized test scores, ELA course grade, and math course grade. Researchers pulled data for each of these variables from the academic year prior to student participation in the BellXcel Summer Program to ensure that the data were not affected by BellXcel participation.

Table 4. Summary of Treatment and Comparison Groups

	Inclusion	**Schools**	**Similarity to Treatment Group**	**Threats to Internal Validity**
Treatment Group	Participates in BellXcel Summer Program	Four Title I OnTrack Greenville treatment schools		
Within-school Comparison Group	Matched to treatment student Did not participate in BellXcel Summer Program	Four Title I OnTrack Greenville treatment schools	Share same school context From similar neighborhoods Equal access to participate in BellXcel Summer Program Access to all outcome measures	Possibility that selection bias is exacerbated by having few high propensity students in comparison group EWRS may encourage all high propensity students to participate in BellXcel, leaving few good matches in comparison group Possibility of spillover effects
Within-district Comparison Group	Matched to treatment student	Four schools identified by the district to serve as comparison schools Schools do not offer the BellXcel Summer Program	Share same district and community context Students who would have participated in the BellXcel Summer Program if it was available to them would be in the control group	Do not share the same school or neighborhood contexts
State Comparison Group	Matched to treatment student	Any Title I middle school in the state Schools do not offer the BellXcel Summer Program	Students who would have participated in the BellXcel Summer Program if it was available to them would be in the control group	Do not share the same school or neighborhood contexts Did not have access to some outcome measures

Researchers estimated a logistic regression using these covariates to produce a predicted probability of receiving treatment for each student separately for the treatment school group, district school group, and state school group. After creating propensity scores, treatment students were matched to comparison students. Like the estimation of the propensity scores, the matching of students occurred independently for each comparison group. Each BellXcel treatment student was matched to five comparison students.

Once the propensity scores were estimated and student matches made, researchers examined the strength of the matches. Ideally, the samples of treatment and district students should be similar to each other, or *balanced*, in terms of the variables used to estimate the propensity score. When samples were not similar, the researchers estimated a new propensity score model using interactions and higher order terms, continuing this process until proper balance was achieved.

At the conclusion of the matching process, researchers ensured that there were no significant differences between the treatment and comparison groups on pre-treatment covariates. Importantly, researchers examined the standardized mean differences and variance ratios between the treatment and comparison groups. Researchers then performed multivariate analyses to provide an estimate of the causal effect of participating in the BellXcel Summer Program.

As detailed below, the matched analyses examined outcomes from both administrative and survey databases. Because of survey nonresponse, the populations of students with complete data were different for the analyses of the outcomes from administrative data and those from the survey data. Therefore, researchers performed separate matching procedures for the outcomes from these different sources. In total, this evaluation includes seven matched comparisons: (1) 2016-17 BellXcel students versus treatment school comparison students on administrative and survey data outcomes, (2) 2016-17 BellXcel students versus district school comparison students on administrative and survey data outcomes, (3) 2016-17 BellXcel students versus state comparison students on administrative data outcomes, (4) 2017-18 BellXcel students versus treatment school comparison students on administrative data outcomes, (5) 2017-18 BellXcel students versus treatment school comparison students on survey data outcomes, (6) 2017-18 BellXcel students versus district school comparison students on administrative data outcomes, and (7) 2017-18 BellXcel students versus district school comparison students on survey data outcomes.

B. Sampling, Measures, and Data Collection

1. Sampling

This evaluation was an analysis of an intensive summer program to diminish summer learning loss in high-poverty schools with a significant population of Hispanic students. The average poverty index for the treatment schools was 83 in academic year 2017-18.[7] The results of this study are generalizable to similar high-poverty schools. The inclusion of the external comparison group, which consisted of students in high-poverty schools with relatively large Hispanic student enrollments across the state of

[7] The SC State Department of Education poverty index is based on Medicaid Enrollment, TANF Enrollment, SNAP Enrollment or Foster Care Services within three years (February 2014 to January 2018) or flagged as migrant or homeless in PowerSchool for academic year 2017-18 (135 Day Census Count).

South Carolina, increased the external validity of this study. Further, this evaluation focused on middle-school students, so the results of the evaluation may not apply to the introduction of the BellXcel Summer program in elementary schools.

The treatment group consisted of all students in the three treatment schools who met the following parameters: (1) the students' parents had agreed to their child participating in the study and (2) the student enrolled in the BellXcel program and attended at least 80% of days offered by the program.

BellXcel aimed to serve 80 students at each of the three middle schools, with an average total enrollment goal of 240 students for each summer. In summer 2016, 233 scholars enrolled in the Summer Program and 199 students completed some portion of the program. In summer 2017, 241 scholars were enrolled and 208 completed some portion of the program. The BellXcel Summer Program is not a mandatory program, so students are not required to attend a specific number of days. Typically, students were dropped from the summer program if they did not participate in the summer program by the beginning of the second week of the six-week program. Students met the treatment definition if they attended at least 80% of the BellXcel Summer Program.

Students in the within-school comparison group also were enrolled in the four treatment schools and these students also had parental permission to participate in the study. If a student ever participated in the BellXcel Summer Program, they were excluded from the within-school comparison group. Final inclusion in the within-school comparison group was determined by the matching process described above.

The population of potential external comparison group students consisted of students in: (1) four other Greenville County Schools middle schools and (2) Title I schools throughout the state of South Carolina. In partnership with researchers, district leaders at Greenville County Schools selected the four within-district schools to serve as comparison schools for the evaluation. There were 19 middle schools in the district during project implementation and the only Title I schools in the district were participating in OnTrack Greenville. Absent other high-poverty middle schools, Greenville County Schools selected the four middle schools with student demographics most similar to OnTrack Greenville schools and a moderate level of student poverty. In academic year 2016-17, 3,398 middle school students attended the district comparison schools, while in academic year 2017-18, 3,568 students attended these schools.[8]

The state comparison students attended Title I schools in districts across South Carolina. Students in Greenville County Schools were excluded from the population of potential state matches, as the presence of OnTrack Greenville programs in the district did not create a "business as usual" comparison and district Greenville County Schools were included in the first external comparison group. In academic year 2016-17, 45,013 middle school students attended a Title I school in South Carolina.

The population of possible external comparison group students included those students who did not have missing data on the variables used in estimating the propensity score. Following the matching procedure described above, external comparison students were matched with treatment students. It is important to remember that this evaluation was a student-level, not a school-level, analysis. However,

[8] Greenville Count Schools Population Statistics for 180th day of attendance, https://www.greenville.k12.sc.us/About/main.asp?titleid=statsarchives

as a means to increase the internal validity of the study, schools were selected such that external comparison students attended somewhat similar schools as the treatment students attended.

For each year of the study, comparison group, and data source, study participants flowed through seven stages in which they either were included or excluded from the study. First, researchers received the roster data for the entire treatment school and district school population. Then, researchers received the roster of students who participated in the intervention from the Sub-Grantee. Researchers only included program students moving forward who met the treatment definition used in the study. The resulting treatment roster was merged with the school population roster. In rare occasions, treatment students did not appear on the school roster and were excluded from the study. Researchers then checked to see which remaining treatment students had complete data for the variables used in propensity score matching. Not all students were successfully matched, as discussed later in this report. Finally, each regression analysis only included students who had data for the outcome variable. Table 5 presents the flow in study participants in academic year 2016-17 for students matched to comparison students in treatment schools using administrative data. Given missing data on the dependent variable, the total number of students included in the final analyses varied somewhat. The values in Table 5 and the other flow charts for the number of students included the final analyses pertain to the most common sample size for the confirmatory analyses. The sample sizes for the other outcomes can be seen in the individual results tables. Similar flow charts for other years, comparison groups, and data sources appear in Appendix C.

Table 5. BellXcel Participant Flow Chart at Treatment Schools AY 2017-18 (Administrative Data)

Study Time-point	Total number students	Number students included	Number students not included	Notes
Treatment Students				
1. Program Roster	230	---	---	
2. Had School Roster Data	230	206	24	
3. Met Treatment Inclusion Criteria	206	109	97	Attended at least 80% of Summer Program
4. Had Full Matching Data	109	106	3	
5. Matched	106	100	6	
6. Included in Main Analyses	100	96	4	
Comparison Students				
1. School Rosters	5,539	---	---	
2. Met Treatment Inclusion Criteria	5,539	505	5034	Treatment School, Didn't participate in program, 6th grade
3. Had Full Matching Data	505	460	45	
4. Matched	460	261	199	Unique students (note: matching was done with replacement)
5. Included in Main Analyses	261	261	0	

Researchers performed a power analysis to assess the relationship between sample size and the minimum detectable effect size (MDES), given a number of assumptions, as part of the Sub-Grantee

Evaluation Plan. In the evaluation presented here, the smallest number of BellXcel students included in the main confirmatory impact analyses was 100. The sample size values included in the SEP power analysis ranged from 120 to 240. The requirement that students attend at least 80% of the BellXcel program to be considered treatment students significantly decreased sample sizes. Post hoc power analyses are uninformative, but one can perform sensitivity analyses that provide the minimum effect size a study could detect given the actual study sample size and a number of assumptions (Perugini, Gallucci, & Costantini, 2018). Using G*Power, a power analysis software program, researchers produced two sensitivity analyses (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). Researchers assumed a one-tailed test with an alpha level of 0.05 and power of 0.80. First, researchers assumed a sample size of 200, which would entail a one-to-one control to treatment ratio. Using this value, researchers estimated that the MDES is d = 0.35. Second, researchers adjusted for the fact that five comparison students were matched to each BellXcel student in this study. Using this five-to-one control to treatment sample size ratio, researchers estimated that the MDES for the study design is 0.20. The small sample sizes in the evaluation limit the study's ability to identify a statistically significant effect of participating in the BellXcel Summer Program. With a larger sample size, some of the small effects detected by this study may have been statistically significant if there had been a larger sample.

Recruitment and Retention

BellXcel staff utilized two primary methods of identifying students who would benefit from participation in the BellXcel Summer program. BellXcel staff worked closely with the district, school principals, and teachers to identify students who would be good candidates for the BellXcel Summer program. In addition, BellXcel used multiple methods to recruit students, including direct mailings, school information sessions with students and parents, flyers and applications sent home in backpacks, phone calls to parents, and neighborhood-based recruitment events.

Keeping the rising middle school students engaged and attending each week was critical to the success of the program. BellXcel used innovative incentives to keep students coming back all summer, including fun, afternoon enrichment sessions, educational field trips, and "Scholar Dollars," which students could use at the Scholar Dollar Store to redeem prizes. In addition, BellXcel worked to consistently encourage parents and guardians to participate in their student's development throughout the summer by inviting them to parent engagement events where students showcased their work and talents. Finally, to help with parent and community engagement, BellXcel recruited bilingual staff, a resource necessary for program success, as the three target middle schools each had a high percentage of Latino students. Students who participated in the BellXcel program but did not meet the minimum participation requirements were not considered "treatment" students for the main treatment-on-the-treated analyses.

In order to maximize participation in the study, the research team employed an opt-out approach to parent consent for the OnTrack Greenville Student Survey administered at treatment and district schools. Parental consent was not needed for the use of student outcome data housed in district and state administrative datasets, as the data were accessed through formal data-sharing agreements.

Attrition and Missing data

Attrition is a challenge for all longitudinal evaluations. This evaluation used a number of means to minimize the effect of attrition. The evaluation team was able to track students who transferred to any other public school in the state. Therefore, these students continued in the study as treatment or

comparison students. Further, since the main outcomes of interest were available in PowerSchool, the state's data system, regardless of the school a student was attending, there were few cases of attrition within comparison groups due to students transferring schools. If study participants were no longer in the state dataset, however, they were considered attrited from the study. Given that consent was a precondition for participating in the treatment group, the evaluators found that consent had little effect on attrition. Parental consent was not needed for the external comparison students, and consent from parents of the within-school comparison group was an opt-out consent, which maximized the number of students participating in this study.

Missing data poses a challenge in every evaluation. The goal of the evaluation team was to limit the amount of missing data, as all adjustments for missing data are suboptimal and impose tradeoffs. Missing data could have occurred in three ways for this study.

First, data could be missing on the receipt of treatment. The BellXcel staff was trained on correctly recording the participation of students in the program. Completing the BellXcel Summer Program was determined by being in the BellXcel system at the beginning and end of the summer program. Students who did not participate in the first week of the program were dropped from the program. The BellXcel staff worked to limit attrition in the summer program.

Second, data could be missing on the main independent variables, including those used for the propensity score analysis. The covariates used in this study, including race, gender, low-income status, and baseline test scores, were all available to the evaluation team in the statewide database for all public school students. Using unique student identification numbers, the evaluation team was able to find the vast majority of students in their dataset regardless of what school(s) a student had previously attended. The state dataset had a few missing cases on the demographic variables, and the evaluators were able to "backfill" any missing data for the permanent demographic variables with datasets from subsequent years. Analyses of some of the secondary outcomes required students to fill out a "baseline" questionnaire. School staff ensured that students completed these questionnaires.

Third, incomplete data could exist in the dependent variable. Incomplete data was a greater issue when examining the survey outcomes. When such cases occurred, the observations were dropped from the analysis (i.e., listwise deletion). Using the U.S. Department of Education's What Works Clearinghouse standards, Puma et al. (2009) recommend case deletion in instances in which post-test or outcome data are missing.

2. Measures and Instruments

Socio-Demographic Variables and Covariates

The following socio-demographic variables were used for propensity score matching and as covariates in impact and outcome analyses: (1) poverty status, (2) race, (3) gender, (4) English proficiency, and (5) disability status.

Poverty status. Student poverty status was measured through free and reduced meal eligibility. As determined by the National Free Lunch Program, students with a family income at or below 130% of the poverty threshold are eligible for free meals, while students with a family income between 130% and 185% of the poverty threshold are eligible for reduced meals (United States Department of Agriculture,

2015). Despite the known limitations for using free and reduced meal eligibility as a proxy for poverty status (Harwell & LeBeau, 2010), this measure could be tracked easily by Greenville County Schools and was readily available for use. Using free and reduced meal eligibility, researchers categorized students as "eligible for free meals or reduced meals" or "not eligible for free or reduced meals." Data on student free and reduced meal eligibility was accessed through district and state administrative records.

Race. There were four categories of student race: "African American," "Caucasian," "Hispanic," and "Other." Student race was accessed through district and state administrative records.

Gender. There were two categories of student gender, "male" and "female." Student gender was captured in district and state administrative records.

English proficiency. Measures of English proficiency consisted of two primary categories, "English language learners" and "non-English language learners." Student English proficiency was accessed through district and state administrative records.

Disability status. The South Carolina Department of Education (SCDE) defines child disability status in accordance with the Individuals with Disabilities Education Act (IDEA) as "having an intellectual disability, a hearing impairment (including deafness), a speech or language impairment, a visual impairment (including blindness), a serious emotional disturbance (referred to in this part as ''emotional disturbance''), an orthopedic impairment, autism, traumatic brain injury, another health impairment, a specific learning disability, deaf-blindness, or multiple disabilities, and who, by reason thereof, needs special education and related services." In the present study, students were categorized as having "no identified disability" or having at least one "identified disability." Student disability status was accessed through district and state administrative records.

Independent Variables

Treatment. The treatment group consisted of all students in the treatment schools who met the following parameters: (1) the student's parent(s)/guardian(s) agreed to their child participating in the study; and (2) the student completed at least 80% of the BellXcel Summer Program. This categorical measure consisted of two groups, "student received treatment" and "student did not receive treatment." Students who participated in less than 80% of the BellXcel program were excluded from both the treatment and comparison sample groups for this analysis.

Primary Impact Variables: Confirmatory

Math and reading course performance. Data from the Measures of Academic Progress (MAP) standardized test measured math and reading course performance. Northwestern Evaluation Association's (NWEA) MAP assessment is a computerized adaptive test for students in grades 2 – 10. Administered up to three times per year, the MAP assessment measures student math and reading achievement and provides immediate results to teachers on student capabilities. Both content area scale scores have shown strong marginal and test-retest reliability in the middle grades and strong concurrent validity when compared to state assessment scale scores (Northwest Evaluation Association, 2004). Further, a confirmatory factor analysis of MAP scale scores across grades and states provided additional support for the construct validity of the instrument (Wang, McCall, Jiao, & Harris, 2012). Researchers accessed MAP assessment data from district administrative records. MAP assessment scores served as a confirmatory impact outcome measure.

There is one important change to the SEP worth noting here. Greenville County Schools stopped administering the MAP assessment in academic year 2017-18, so researchers did not have post-program MAP assessment data for summer 2017.

Primary Impact Variables: Exploratory

Behavioral referrals. The primary measure of student behavior was the total *number of behavioral referrals* per student. The number of behavioral referrals was available in the district dataset. The state dataset included the number of discipline incidents, so that is the outcome used for the state analyses. Previous research has found office behavioral referrals to be a meaningful source of data for designing and evaluating behavior interventions (Putnam, Luiselli, Handler, & Jefferson, 2003; Sugai, Sprague, Horner, & Walker, 2000). In addition, researchers used a dichotomous categorical variable for student behavioral referrals, with students either having "no behavioral referrals" or "one or more (any) behavioral referral(s)." Researchers accessed behavioral referral data from state and district administrative records.

In-school suspensions. This measure included the total *number of hours of in-school suspension* served by the student within the academic year. The number of hours of in-school suspension was available in the district dataset. The state dataset included the number of in-school suspensions, so that is the outcome used for the state analyses. In addition, researchers used a dichotomous categorical variable for student in-school suspension, with students either having "no in-school suspensions" or "one or more in-school suspensions." Researchers accessed in-school suspension data from state and district administrative records.

Out-of-school suspensions. This measure included the total *number of days of out-of-school suspension* served by the student within the academic year. For the state analyses, the total number of out-of-school suspensions was used. In addition, researchers slightly modified the SEP to include a dichotomous categorical variable for student out-of-school suspension in the analysis, with students either having "no out-of-school suspensions" or "one or more out-of-school suspensions." Researchers accessed out-of-school suspension data from state and district administrative records.

Average daily attendance. This measure of school attendance was a calculation of the number of days of school attended divided by the number of days of school enrolled. Researchers accessed average daily attendance data from state and district administrative records.

Chronic absenteeism. Chronic absenteeism occurs when a student is absent more than 10% of the academic year, including both excused and unexcused absences. At the school level, the number of chronically absent students is often more telling than the average daily attendance rate, as a school could boast an acceptable attendance rate and still have a large number of students chronically absent given the distribution of absences (Balfanz & Byrnes, 2012). Though not commonly tracked by school districts, this measure is growing in popularity, especially due to its utility as an early warning indicator (Balfanz, Herzog, & Iver, 2007). In the present study, chronic absenteeism is a dichotomous categorical variable with students either "chronically absent" or "not chronically absent." Researchers accessed chronic absenteeism data from the state and district administrative records.

Secondary Outcome Variables: Exploratory

Students increase self-confidence. This construct was measured by two scales, the Academic Perseverance scale and the Academic Self-Confidence scale (Rockman et al, 2013).

The six-item Academic Perseverance scale measured having a hopeful outlook on studying and completing schoolwork and included items such as "I keep doing schoolwork even when it is hard" and "When I study, I set goals for myself." Items were scored on a four-point Likert-type scale ranging from "not true" to "true." This scale demonstrated strong internal reliability in prior studies ($\alpha = 0.79$) and in the current study ($\alpha = 0.76$).

The six-item Academic Self-Confidence scale measured the ease with which students felt they were learning and included items such as "Homework is easy for me" and "I understand what we are learning in school as much as my friends." Items were scored on a four-point, Likert-type scale ranging from "not true" to "true." This scale demonstrated strong internal reliability in prior studies ($\alpha = 0.75$) and in the current study ($\alpha = 0.78$).

Students improve their attitude toward learning. This construct was measured by a modified version of the Valuing Education scale (Corrin et al., 2015). Originally, this was a six-item scale measuring a student's beliefs about the importance of school and included items such as "My education will be valuable in getting the job I want" and "Being a good student is important to me." Researchers added one additional item on the importance of attending college to make this a seven-item scale. Items were scored on a four-point, Likert-type scale ranging from "not true" to "true." The original scale demonstrated strong internal reliability in prior studies ($\alpha = 0.79$). In the present study, the modified scale was tested using exploratory factor analysis with oblimin rotation and was found to have a unidimensional factor structure, as predicted, with strong internal reliability ($\alpha = 0.84$). This construct was measured by BellXcel administrative data tracking parent attendance at BellXcel family engagement activities. In addition, this construct was measured by a single-item survey question on the BellXcel Parent Survey: "I am more involved in my child's learning." This question had three response categories, "agree," "disagree," and "not sure."

Students' increased reading and math skills. Student reading and math skills were measured using the STAR Assessment. A computerized adaptive test, the STAR assessment helps educators evaluate scholar improvement, differentiate instruction, identify struggling scholars, and group scholars based on their academic areas of need. Researchers accessed STAR Assessment data from BellXcel program records.

Parents increase engagement in child's education. Researchers measured parent engagement through a single-item survey question on the BellXcel Parent Survey: "I am more involved in my child's learning." This question has three response categories, "agree," "disagree," and "not sure."

Researchers intended to track attendance data from parent engagement events, but BellXcel program records did not provide enough information to allow researchers to examine the data in-depth. Thus, this construct was only measured by the single survey item described above.

Additional Exploratory Outcomes

In order to reduce data collection activities across multiple OnTrack Greenville studies, researchers included additional outcome measures on survey instruments. The following outcomes are not part of the BellXcel Summer Program logic model, but data were available and included in exploratory analyses.

Students build relationships with caring adults. This construct was measured by two scales, Relationships with Caring Adults (Corrin et al., 2015) and Positive Relationships with Teachers (Corrin, Sepanik, Rosen, & Shane, 2016). The six-item Relationships with Caring Adults scale measured the extent to which students related to school personnel and included items such as "At my school there is an adult who really cares about me" and "At my school there is an adult who always wants me to do my best." Items were scored on a four-point, Likert-type scale ranging from "not true" to "true." This scale demonstrated strong internal reliability in prior studies ($\alpha = 0.89$) and in the current study ($\alpha = 0.89$).

The eight-item Positive Relationships with Teachers scale measured the perceived relationships of students with their teachers and classmates and included items such as "Students at my school get along well with teachers" and "My teachers really listen to what I have to say." Items were scored on a four-point, Likert-type scale ranging from "not true" to "true." This scale demonstrated strong internal reliability in prior studies ($\alpha = 0.86$) and in the current study ($\alpha = 0.85$).

Students are more engaged at school. This construct was measured using the School Engagement scale (Rockman et al, 2013) and the School Belonging scale (Corrin et al., 2015).

The four-item School Engagement scale measured the degree to which a student felt connected to his or her school and education (Rockman et al, 2013). It included items such as "I like school" and "I participate a lot in class." Items were scored on a four-point, Likert-type scale ranging from "not true" to "true." This scale demonstrated strong internal reliability in prior studies ($\alpha = 0.68$) and in the present student ($\alpha = 0.74$).

The five-item School Belonging scale measured the extent to which a student felt accepted and supported within the school environment (Corrin et al., 2015). It included items such as "I feel close to people at my school" and "I feel like I am a part of my school." Items were scored on a four-point, Likert-type scale ranging from "not true" to "true." This scale demonstrated strong internal reliability in prior studies ($\alpha = 0.84$) and in the present student ($\alpha = 0.83$).

Frequency of summer activities. A series of seven individual survey items measured the frequency with which students participated in a variety of activities over their summer vacation. Sample items included "How often did you go to the library?" and "How often did you play math games or solve math problems?" These items were scored on a five-point Likert-type scale ranging from "never" to "very often (nearly every day)."

3. Data Collection Activities

The measures used in the impact study for propensity score matching, covariates, independent variables, and primary impact variables were collected routinely by Greenville County Schools using the PowerSchool data management platform. The impact study drew on student data from both Greenville

County Schools and the South Carolina Department of Education. Researchers collected data to explore the secondary outcomes via the administration of electronic student surveys.

Student data from the South Carolina Department of Education. The South Carolina Department of Education (SCDE) mandates the use of PowerSchool and provides districts technical manuals and support to improve the internal reliability of data collected. The SCDE routinely collects and aggregates data from all districts and houses it in the South Carolina Education Data System (SCEDS). The Riley Institute currently has a Memorandum of Agreement with SCDE that dictates the terms and conditions of the transfer of PowerSchool data from SCDE to the Riley Institute, including provisions for maintaining, protecting, and destroying datasets. The inclusion of these data allowed researchers to examine a group of comparison students from schools across the state of South Carolina.

Student data from Greenville County Schools. A research and data sharing agreement with Greenville County Schools also provided administrative data for the study. This data sharing agreement describes (1) the research and information usage terms and conditions; and (2) the purpose and design of the study, including type(s) of data requested, data collection schedule, plan for reviewing and sharing results, and methods of securing and destroying data.

OnTrack Greenville Student Survey. The research team administered the OnTrack Greenville Student Survey to collect data for secondary outcome measurement. Teachers administered the survey to students electronically in October and May of academic years 2016-17 and 2017-18. The Research Team obtained passive parental permission by sending home an opt-out letter at the beginning of the academic year. Opt-out consent was sufficient, as the survey did not include any identifiable information.

In order to link the survey data to the PowerSchool dataset, researchers created a unique survey ID number for each student who was not opted-out by their parents. Researchers maintained a separate database that linked the survey ID numbers with each student's PowerSchool ID number. In preparation for survey administration, the researchers created individual notecards for each student, which included the student's name and unique survey ID. Researchers organized the notecards in packets by school and teacher and distributed the packets to the schools in-person, along with survey administration instructions, before the survey administration window opened. Teachers passed out the notecards to each student whose parents did not opt them out of the survey (teachers were provided a list of those students who had been opted out) and provided oral instructions on how to complete the survey. To begin, students navigated to the electronic survey on a computer and entered their unique survey ID to link their answers to their PowerSchool data. Students then were given permission to opt themselves out of the survey if they decided they did not want to take it. Survey completion took between 10 and 20 minutes, with an average student completion time of 12 minutes. Following administration, the teacher collected all survey ID cards and returned them to their survey packet. Researchers then destroyed all of the ID cards. The OnTrack Greenville Student Survey is located in Appendix D.

Information on the response rate of the OnTrack Greenville Student Survey appears in Tables 6 and 7. All students attending treatment and district schools within Greenville County were invited to participate in the pre- and post-survey. The parent opt-out rate was slightly higher at treatment schools than comparison schools, largely due to differences in distributing the passive parental permission form. At treatment schools in academic year 2016-17, the permission form was sent home with students at the beginning of the academic year along with other first-day-of-school forms. Many of these forms had to be returned with a parent signature, so researchers suspect that many parents signed and returned

the opt-out form to decline their child's participation in the study without reading the form completely. At comparison schools, however, the passive parental permission form was sent home with students a few weeks after the start of the academic year and fewer forms were returned. For that suspected reason, the rate of student opt-out was 5% at treatment schools and 1 – 2% at comparison schools in the 2016-17 academic year.

Students who were absent the day of survey administration did not have an opportunity to complete the survey at a later date. In addition, individual schools decided not to offer the survey to students with severe learning and/or intellectual disabilities. When cleaning the data, researchers used list-wise deletion to eliminate cases with missing data. In academic year 2016-17, after excluding these students and cases, the total percentage of valid survey responses for each treatment group at each survey administration ranged from 65 – 72% at treatment schools and was 79% at comparison schools in the 2016-17 academic year. Treatment school students took an average of 13 minutes to complete the OnTrack Greenville Student Survey, while district school students took an average of 11 minutes to complete the survey.

Table 6. Response Rate of OnTrack Greenville Student Survey AY 2016-17

	Pre-Survey Fall Semester 2016				Post-Survey Spring Semester 2017			
	Treatment Schools (n =4)		Comparison Schools (n =4)		Treatment Schools (n =4)		Comparison Schools (n =4)	
Total # Invited to Participate	1,921	100%	3,369	100%	1,886	100%	3,368	100%
Parent Opt-Outs	212	11%	91	3%	195	10%	208[1]	6%
Student Opt-Outs	89	5%	47	1%	90	5%	65	2%
Excluded Cases[2]	238	12%	560	16%	384	20%	445	13%
Total # Valid Survey Responses	1,382	72%	2,671	79%	1,217	65%	2,650	79%
Average Completion Time	13 minutes		11 minutes		13 minutes		11 minutes	

[1]Mid-year transfer students who never received a parental permission form were treated as parent opt-outs at comparison schools.
[2]Excluded Cases includes students who were absent the day of the survey, duplicate survey starts, incomplete survey responses, etc.

For the 2017-18 OnTrack Greenville Student Survey, the total percentage of valid survey responses for each treatment group at each survey administration ranged from 70 – 75% at treatment schools and from 71 – 81% at comparison schools in the 2017-18 academic year. Treatment students took an average of 12 minutes to complete the OnTrack Greenville Student Survey in the 2017-18 academic year, while district students took an average of 10 minutes.

Table 7. Response Rate of OnTrack Greenville Student Survey AY 2017-18

	Pre-Survey Fall Semester 2017				Post-Survey Spring Semester 2018			
	Treatment Schools (n = 4)		Comparison Schools (n = 4)		Treatment Schools (n = 4)		Comparison Schools (n = 4)	
Total # Invited to Participate	2,040	100%	3,692	100%	2,069	100%	3,498	100%
Parent Opt-Outs	45	2.2%	82	2.2%	43	2.1%	292[1]	8.3%
Student Opt-Outs	127	6.2%	91	2.5%	168	8.1%	117	3.3%
Excluded Cases[2]	348	17.0%	536	14.5%	397	19.2%	599	17.1%
Total # Valid Survey Responses	1,520	74.5%	2,983	80.8%	1,461	70.1%	2,490	71.1%
Average Completion Time	11m 45s		10m 45s		12m 15s		9m 30s	

[1]Mid-year transfer students who never received a parental permission form were treated as parent opt-outs at comparison schools.
[2]Excluded Cases includes students who were absent the day of the survey, duplicate survey starts, incomplete survey responses, etc.

BellXcel Parent Survey. As part of the program's internal performance management monitoring, BellXcel program staff administered the Parent Survey to parents at two of the three sites. An anonymous survey, parents completed the questionnaire on paper. BellXcel administered the survey in two ways: (1) handing out the survey to parents in person at family events; and (2) sending the survey home with students and asking parents to return the completed survey with their student. The survey was available in English and Spanish and required between 5 and 10 minutes to complete. In summer 2016, 43 parents completed the survey, while 98 parents completed the survey in summer 2017.

III. Implementation Study

A. Implementation Study Design

The Riley Institute at Furman University contracted with RTI International to conduct an implementation evaluation of the BellXcel Summer Program in three Title I middle schools in the summers of 2016 and 2017.

Researchers approached the implementation study from utilization-focused and developmental evaluation perspectives and were committed to gathering and presenting information that could be used by BellXcel and its Greenville partners to improve implementation over time, while also recognizing areas of strengths. Researchers also were sensitive about conducting the study efficiently and not overburdening BellXcel staff, scholars, and parents. The BellXcel logic model has been studied by other research institutions and has comprehensive internal monitoring and evaluation tools, including surveys of BellXcel teachers, scholars, and parents. Rather than adding new sets of data collection processes, researchers leveraged existing tools and processes, requesting some modifications and a small set of new indicators as needed.

B. Data Collection Methods

Scholar Survey

Between both years of implementation, a total of 236 scholars across all three sites completed at least part of the survey. Administrative data provided by BellXcel and included in Table 9 below reports that 407 scholars completed the program, yielding an overall response rate of 79%. However, administrative issues at School 3 both years of implementation, resulted in lower response rates for School 3. Broken down by site, 107 out of 147 (73%) scholars responded at School 4, 76 out of 113 (67%) responded at School 1, and 53 out of 147 (36%) responded at School 3. Results did not vary substantially across sites, therefore, in this report the results are presented in the aggregate across all three sites for most survey questions. The BellXcel Scholar Survey appears in Appendix E.

Staff Survey

At the end of each program, BellXcel administered an online survey to staff at all sites in operation during the summers of 2016 and 2017. Responses were analyzed for the staff survey administered at the three Greenville sites—Schools 1, 3, and 4. As with the scholar survey, researchers reviewed the standard staff survey and requested a few additional items be added to the versions administered to the Greenville sites.

A total of 44 BellXcel staff completed the staff survey, resulting in an approximate survey response rate of 58%.[9] Given the low numbers and the desire to preserve anonymity, results from the staff survey are not broken out by site but are instead presented only in the aggregate.

[9] The denominator for the response rate includes the 76 BellXcel staff members who were classified as academic teachers, enrichment teachers, academic and enrichment teachers, assistant teachers, instructional coaches, program managers, and program assistants. It excludes substitute teachers, who were not asked to complete the survey.

Some 30% of BellXcel staff who completed the staff survey were assistant teachers, 30% were academic teachers, 15% were program/site managers, and 10% were assistant program/site managers. Five percent each were dual-role teachers, enrichment teachers, and association/partnership program leads. Substitute teachers were not asked to complete the survey. The BellXcel Staff Survey appears in Appendix F.

Parent Survey

BellXcel asks parents of BellXcel scholars to complete a paper survey at the end of the program. Researchers reviewed the standard parent surveys and requested a few additional items for the 2016 summer program in Greenville, and these items remained on the 2017 parent survey. A total of 144 parents responded, 51 from School 1, 44 from School 3, and 49 from School 4. The number of respondents was substantially higher in 2017 (101) compared with summer 2016 when only 43 parents completed surveys. The parent survey was available in English and Spanish. The evaluation team noted a potential translation error on the Spanish surveys, therefore, data for the questions soliciting feedback about the orientation and family week event were limited only to those who indicated they attended and provided feedback.[10] No demographic information about parents was collected. The BellXcel Parent Survey appears in Appendix G.

Observations, Interviews, and Focus Groups

Each summer, two researchers conducted site visits in Greenville, spending a full day at each of the three middle schools both years of implementation. During the site visits, researchers observed morning community time, academic (both ELA and math) classes, and enrichment classes. At all three sites, researchers conducted interviews with the program managers and instructional coaches. In addition, at each site, researchers conducted a focus group with teachers. Qualitative findings from these observations, interviews, and focus groups are included in the analysis when relevant. The Program Manager Interview Protocol is located in Appendix H and the Teacher Focus Group is located in Appendix I.

Administrative Data from BellXcel

While developing the implementation evaluation plan, researchers worked with BellXcel's Director of Evaluation to clarify the indicators for which researchers would need data, which could not be gathered through the surveys or site visit data collections. This included information such as BellXcel staff training completion rates, scholar attendance rates, and bus usage rates. In the fall, BellXcel transferred the administrative data to researchers.

[10] In some cases, parents indicated they did not attend the event but provided feedback. This may have been due to a translation error on the survey. These data were excluded from analyses.

C. Program Inputs and Activities

RQ12. Were program components (such as staffing, curriculum, data systems, provision of transportation, and community engagement events) implemented as intended at each school?

In both 2016 and 2017, one hundred percent of BellXcel staff at the three sites completed the in-person BellXcel training. The first year of implementation saw 94% of staff at School 4, 92% at School 3, and 73% at School 1 completed the additional mandatory online training.[11] In summer 2017, the online training was optional, with content covered during the in-person training instead. As a result, a lower percentage of teachers took the additional online training in the second year of implementation, with 87% of staff at School 3, 64% at School 4, and 50% at School 1 completing the online training. During both years of implementation, teachers (including lead teachers, assistant teachers, and instructional coaches) were overall satisfied with the extent to which the training prepared them for their roles.

A key input of the BellXcel logic model is a curriculum aligned with the Common Core. BellXcel teachers reported implementing the BellXcel curriculum in both ELA and math, although many teachers indicated they modified or supplemented the curriculum to meet the needs of their scholars, many of whom were performing substantially below grade level. During both years of implementation, veteran teachers noted improvements in the BellXcel math curriculum by broadening the focus beyond fractions.

In addition to ELA and math, all three sites implemented an enrichment curriculum. During the site visit, the researchers observed enrichment classes in coding, STEM courses, nutrition, dance, social skills, technology, cooking, art, and karate. In every case, scholars had access to appropriate and interesting materials and appeared engaged and joyful. In most of these classes, teachers drew connections to careers that aligned to the activity (e.g., chef and chemist for cooking, engineer and graphic designer for art).

Community time, another curricular element designed to start the day with a fun, interactive, and structured activity, was executed well across the three sites in 2017. This is an improvement from the first year of implementation, when only one site appeared to implement community time with fidelity.

Teachers generally reported positive views on behavior management, with 90% of teachers reporting that they rarely had behavior problems with their scholars in 2016. Additionally, 100% of teachers in 2017 agreed or strongly agreed that the behavior management system allowed for fair and respectful treatment of scholars and that it allowed for scholars to learn self-management. Across both years, scholars were slightly less positive about scholar behavior and views varied by site and year. For example, 40% of School 1 scholars reported that scholars behaved well almost always, compared to 6% at School 4. Comparatively, 38% of School 4 scholars reported that scholars almost always behaved well, compared with 35% at School 1 and 24% at School 3.

RQ13. *Did the quality of the program reflect the intended design?*

In general, researchers found a relatively high degree of alignment between the intended and actual implementation of BellXcel at each of the three Greenville middle schools. Implementation in regard to teachers showed that 100% of teachers completed the in-person BellXcel training, whereas a smaller percentage of teachers completed at least 90% of the online course (what was required to have that be

[11] Completion is defined as finishing at least 90% of the course.

considered complete). Teachers implemented elements of the program, including the enrichment curriculum, community time, and field trips, with more alignment to the program design in 2017 than in 2016.

In regard to scholars, they were provided the proper dosage of the BellXcel program. Over both years of implementation, the BellXcel Summer Program at the three sites in Greenville operated for slightly more than 6 weeks. The daily schedule at each site indicated that scholars generally received academic instruction for 3 hours (8:30 to 11:30), Monday through Thursday, and enrichment for 2 hours (12:20 to 2:20), Monday through Wednesday. On Thursdays, scholars generally attended a field trip in the afternoon. 64% in 2016 and 62% in 2017 of students attended at least 80% of program days.

Parent engagement events are an important part of the BellXcel Summer Program. In Greenville, each site conducted a parent orientation, an open house, a family literacy workshop, and a closing ceremony. The logic model specifies a goal of 70% of parents attending parent engagement events. No site reached 70% parent participation at any one event in either 2016 or 2017. Parent survey respondents who attended the parent engagement events generally had positive feedback about both events, particularly the family literacy workshop.

At least 70 percent of parents from both years of implementation agreed that events helped them to understand their scholars' experience in the BellXcel program, were well planned and had a clear agenda, and provided the opportunity for them to connect with BellXcel teachers and staff.

RQ14. *What aspects of the design were modified at each school, if any, and why were they modified? What were the barriers to implementation, if any?*

In general, researchers found a relatively high degree of alignment between the intended and actual implementation of BellXcel at each of the three Greenville middle schools. Teachers in both years of implementation expressed difficulties implementing the math curriculum for a variety of reasons and modified the curriculum to match scholar's needs.

Regarding behavior management techniques, less than 50% of teachers used the BellXcel behavior management system. Regardless of the strategy used, all staff either agreed or strongly agreed that when scholars misbehaved, staff members were comfortable managing the situation.

There were no systematic barriers to implementation. The most common reasons scholars gave for missing one or more days of the BellXcel Summer Program during both years of implementation were sickness, family vacation, and doctor's appointment or other appointment. Scholars were asked to write in other reasons they missed 1 or more days; responses included "because my uncle was very sick," "family emergency," "I overslept," "one day no one was at home," and "I didn't know [BellXcel] was starting."

In 2016, about 15% of scholars indicated that they missed one or more days because they had trouble finding a ride to the program. By site, about 23% of scholars at School 1 indicated that they missed a day for this reason, compared to 10% at School 4.

D. Teacher Activities and Outputs

RQ15. *Did 100 percent of teachers, teaching assistants, and enrichment assistants receive the proper dosage of training?*

A key output in the BellXcel logic model is that 100% of BellXcel staff complete BellXcel training, including both the "BellXcel University" e-learning platform and the in-person training prior to the start of the program. Data on training completion came from BellXcel's administrative records.

Overall, 100% of all teachers, program managers, program assistants, and instructional coaches across all sites completed the in-person BellXcel training during both years of implementation. To be considered as having completed the online training, staff must have finished at least 90% of the course.

Table 8. Percent of staff who completed at least 90% of the online course, by site and year

Site	Summer 2016	Summer 2017
School 1	90%	50%
School 3	94%	87%
School 4	92%	64%

Note: The online training course was a mandatory part of training in 2016 and optional in 2017.

In 2016, the online training course was mandatory and 94% of School 4 staff and 92% of School 3 staff completed at least 90% of the online course. Completion was somewhat lower at School 1, where 73% of staff completed at least 90% of the online course. In 2017, the online training course was optional, with core content covered at the in-person training. As a result, a smaller percentage of teachers completed the online training in 2017. For example, 87% of School 3 staff and 64% of School 4 staff completed the online course. Completion was lower at School 1 where 50% of staff completed the online course.

RQ16. *To what extent did the training provide teachers and assistants with the necessary skills and knowledge they needed to implement the program? What additional knowledge or skills did teachers and assistants require, if any? To what extent were teachers able to find support to fill in knowledge or skill gaps?*

In 2016, survey responses indicated that BellXcel teachers[12] were generally satisfied with the training they received. 95 percent agreed or strongly agreed that the classroom training was high quality and 90 percent agreed or strongly agreed that the online training was high quality. The vast majority of teachers (90 percent or more) also agreed or strongly agreed that after training they felt prepared to be role models for scholars, to manage behavior, and to work as a collaborative team. However, almost a quarter disagreed or strongly disagreed that the online training was user-friendly and structured in a way that was easy to understand.

In 2017, survey responses indicated that BellXcel teachers[13] were generally satisfied with the training they received, as the majority felt either somewhat or really well prepared for BellXcel-related tasks. For example, 100% indicated they were somewhat prepared to be a role model for their scholars (93% felt

[12] For the purposes of these sections, "teachers" include academic teachers, enrichment teachers, assistant teachers, and instructional coaches.

[13] In this section, "teachers" include academic teachers, enrichment teachers, and assistant teachers,

really well prepared), and 93% felt somewhat prepared to manage behavior in their classrooms. Slightly fewer (though still most) teachers felt somewhat prepared to utilize assessment data to impact scholars' academic development (86%), implement math curriculum in their classrooms or implement growth mindset activities (79% each), and implement ELA curriculum in their classrooms (72%). The vast majority of teachers (86%) also reported that they felt somewhat prepared to work as a collaborative team with their summer colleagues.

Qualitative data collected through the teacher focus groups were relatively consistent with the survey data, with the focus group respondents generally reporting having received sufficient training, although some of the new teachers believed it was a bit rushed. One teacher suggested different trainings for returning and new teachers, with the former receiving a condensed refresher. A few respondents suggested that the online course was helpful as a new teacher or assistant but was not necessary for returning staff. The teachers appreciated the planning time they had within ELA and math teaching groups, including with teachers from the other sites and the instructional coaches.

RQ17. *To what extent did teachers implement the curriculum as intended? What challenges emerged as teachers implemented the curriculum?*

Academic curriculum

A key input of the BellXcel logic model is a curriculum aligned with the Common Core. In focus groups with BellXcel teachers and interviews with BellXcel program managers and instructional coaches, staff reported implementing the BellXcel curriculum in both ELA and math and that the curriculum was, in general, of high quality. Nevertheless, many staff members reported modifying or supplementing the curriculum to meet the needs of their scholars. For example, math teachers at School 4 reported they had substantial numbers of scholars who had higher-level math skills than scholars the BellXcel curriculum served, and the teachers had to supplement the curriculum with other materials. Teachers at School 1 reported a different problem: many of their scholars came to BellXcel far below grade level, and the BellXcel material did not meet their needs. Similar comments were shared by teachers across the three sites with regard to ELA: teachers needed more materials appropriate for scholars who were reading far below grade level. Teachers at all three sites also noted that the ELA pacing guides, while helpful in theory, were unrealistic. Some teachers trusted their professional judgment and did not worry about not being able to follow the pacing guides. Others were more troubled by the discrepancy between the time specified in the guides and how much time tasks actually took.

Teachers, program managers, and instructional coaches who were veterans of BellXcel noted improved math curriculum from previous years. Specifically, they appreciated that a more diverse set of concepts were covered beyond fractions, which had previously been the primary focus. They also applauded the incorporation of *Dyno Math*, a math magazine that scholars found engaging.

Teachers at one site had very strong and negative reactions to what they explained was a new approach to and template for writing lesson plans. They were presented with the template during the training and believed it was very detailed and not well explained. When they used the template to write their first lesson plans, they were very confused and said it took 4 hours. They thought writing plans and following steps was emphasized over just doing what was necessary to support scholars with the lessons. The instructional coach at that site voiced a similarly strong critique of the lesson plan process. Researchers visited this site on the first day so afterward listened for similar concerns at the other two sites. When

such concerns were not brought up, the focus group leader asked about the new lesson planning template. Teachers and the instructional coaches at the other two sites were nonplussed. They did not find the lesson planning template or process to be overly burdensome or require much time.

Enrichment Curriculum

In addition to ELA and math, all three sites implemented an enrichment curriculum. The BellXcel enrichment curriculum is designed to expose scholars to cultural and recreational activities they are not likely to have in school and to develop skills that will open up new career and educational opportunities and spark new interests. During the 2016 site visit, researchers observed varying degrees of quality of implementation, with some courses not having appropriate materials and infrastructure. For example, in one cooking class, the teacher, without access to a kitchen, was limited to reviewing a recipe and discussing fractions in the context of measurement while scholars sat at their desks. The enrichment courses observed during the 2017 site visits all appeared well organized with sufficient materials and enthusiastic teachers who proactively drew connections for scholars between the activities they were working on, topics they could study in college, and careers they could pursue. For example, scholars in an art class were developing a model in pairs from an assortment of seemingly random materials. The teachers introduced engineering concepts and encouraged scholars to test different designs.

Other Curricular Elements

Community time, typically combined with breakfast, is a required element of the BellXcel program and lasts at least 20 minutes per day. The purpose of community time is to build site-wide group identity and relationships; support scholar development of social skills, cooperation, empathy, self-control, and self-confidence; and start the day on a positive note with a fun interactive structured activity. For example, the BellXcel handbook indicates that community time might consist of call and response greeting, an activity, sharing, and news and announcements.

During the summer 2016 site visits, researchers did not observe community time being implemented consistently across sites. In summer 2017, however, researchers found that community time was being used as intended—with scholars playing team-building games and teachers having positive conversations with scholars as they ate breakfast.

Another important feature of the BellXcel program is field trips, which are intended to expose scholars to resources in the community, provide scholars with opportunities they might not otherwise have, and enhance fun. All three sites incorporated field trips into the summer program, and teachers across sites had high praise for the field trips. For example, during the week researchers visited, scholars were preparing to visit Furman University.

In addition to field trips, special events were held at the school sites to support extra-curricular learning and also a spirit of celebration, which seemed to the observers to be important in order to keep scholars engaged and enthusiastic. The day that researchers were at School 1, four players from a local professional baseball minor league team visited and talked with scholars about their college and career journeys thus far. At School 3, all scholars were gathered for a demonstration and practice in using combinations locks, which they will use when they are assigned lockers at the start of 6th grade. The scholars were very excited, took the practice time seriously, and were clearly proud when they successfully unlocked their lock, often after a number of attempts. At School 4, the program manager

had arranged for a surprise for the scholars towards the end of their field day activities: a food truck serving free icy treats.
Finally, staff at all sites reported incorporating a college-going culture into the curriculum. All classrooms displayed college posters, and at one site classrooms were named for the colleges attended by the teachers. In both academic and enrichment classes, teachers used language that assumed all scholars would attend college (e.g., "when you go to college, you're going to need to know how to manage your time.")

Behavior Management

BellXcel provides teachers and staff with a behavior management system. On the 2016 staff survey, 45 percent reported using BellXcel's behavior management system, while 30 percent reported using a school-based system and the other 25 percent reported using a different system (Exhibit 8). Similarly, in 2017, 50% of staff reported using a school-based behavior management system, 25% reported using BellXcel's behavior management system, and 25% reported using a different system. Two respondents indicated they used something similar to the school-based system but "focused on student engagement, encouragement, and rewards for positive behavior."

In 2016, 90 percent of teachers agreed or strongly agreed that they rarely had behavior problems with their scholars, and 100 percent of teachers reported they felt comfortable dealing with scholar behavior themselves. However, 30 percent agreed or strongly agreed that they spend a lot of time trying to get scholars to settle down and stop talking.

Regardless of the strategy used, 100% of teacher survey respondents in 2017 either agreed or strongly agreed that their behavior management systems allowed for fair and respectful treatment of scholars and allowed for scholars to learn self-management. Further, all staff either agreed or strongly agreed that when scholars misbehaved, staff members were comfortable managing the situation and that scholars knew and understood the consequences if they acted out.

Scholars reported a somewhat more negative perception of scholar behavior than did staff. In 2016, about 16 percent of scholars reported that scholars behaved well in class only once in a while or almost never, while 28 percent reported that scholars behaved well sometimes. In comparison to 2017, about 9% of scholars reported that scholars behaved well in class almost never or once in a while, while 30% reported that scholars behaved well sometimes. School 1 scholars reported the most positive behavior among their peers: 72% believed that other scholars often or almost always behaved well in class, compared with 59% and 53% for School 4 and School 3 scholars, respectively. However, this discrepancy was not as pronounced as in summer of 2016, when the gap was 32 percentage points between School 1 and School 4 (insufficient sample size in 2016 for School 3 precluded reporting its results separately. The results for the total set of scholars revealed a larger percentage of 2017 scholars believed their peers behaved well in class often or almost always compared with 2016 scholars (61% compared with 47%).

RQ18. *What suggestions do teachers have for improvement of the various program components as implemented?*

In summer 2017, survey data and focus group results indicated that teachers had mostly positive perceptions of program leadership. For example, on the survey, nearly 80% of teachers indicated they felt somewhat supported or very supported by their site leaders in curriculum implementation,

developing skills in managing scholar behavior, developing instructional plans, and developing and adjusting instructional delivery. In the focus groups, teachers praised the program managers and instructional coaches at their sites and indicated that the program was well run and organized. Teachers also acknowledged how accessible the Greenville BellXcel Director was and appreciated that he could interface with BellXcel National staff when needed to answer questions.

Teaching staff members were asked to provide feedback about their interactions and communications with site leadership. Responses to questions about communication from site leadership varied more than questions about support. For example, over 90% of staff reported that policies for program staff and the payroll process were well communicated or really well communicated (100% and 93%, respectively. In contrast, only 71% of staff indicated that teaching, mentoring, and/or child development strategies were well communicated or really well communicated, and only 64% reported that site information about upcoming events was well communicated or really well communicated.

Finally, teachers had very positive overall feedback about site leadership. Over three-quarters of teachers either agreed or strongly agreed that program leaders were highly visible around the site (100%), that program/site leaders promoted teamwork at their sites (93%), and that they received the tools and resources they needed to their job well (86%).

In sum, teachers were overall very happy with the program as implemented. However, there were some suggestions for improvements, including the following:

- Have different training tracks for new and returning BellXcel teachers.
- Simplify the lesson plan template and process.
- Make the pacing guides more realistic and less structured, or clearly acknowledge that teachers can use their professional judgment to make adjustments.
- Focus more on helping teachers differentiate instruction for the wide range of scholar needs.

D. Scholar Attributes, Dosage, Activities, and Outputs

RQ19. *How were scholars recruited at each site and to what extent was EWRS used to identify scholars? What percentage of scholars who participated in BellXcel at each site meet the original target characteristics specified in the model (at or below the 40th percentile in English Language Arts or mathematics)? Which, if any, additional target characteristics were added during the recruitment phase? If additional criteria were added, what percentage of scholars met each of the target characteristics? What recommendation do district and BellXcel staff have for improving recruitment next year?*

The Early Warning and Response System, which is operated by the district and flags students who are going off course in course performance, attendance, or behavior, has not yet been used to identify scholars to invite to BellXcel, although it may be used in future years. Instead, MAP scaled and percentile scores were used to identify fifth-grade students in the district who were at or below the 60th percentile in ELA or math. Although the BellXcel National program describes its model as being most appropriate for students at the 40th percentile or below, BellXcel Greenville focused its recruitment for summer 2017 on students at the 60th percentile or below due to the low recruitment numbers from the previous year, when BellXcel and district staff attempted to recruit students only at the 40th percentile or below. In 2017, BellXcel hired a recruitment specialist earlier in the spring semester than they did in

2016. The pace of registration improved across the 2 years. Whereas in 2016 enrollment targets had not been met by May, in 2017 they were met, and there were waiting lists.

RQ20. *Were scholars provided the proper dosage of the program (i.e., six weeks operating six and a half hours per day, four days per week, with three hours of academic instruction and two hours of enrichment)?*

In general, scholars were provided the proper dosage of the BellXcel program. Over both years of implementation, the BellXcel Summer Program at the three sites in Greenville operated for slightly more than 6 weeks. The daily schedule at each site indicated that scholars generally received academic instruction for 3 hours (8:30 to 11:30), Monday through Thursday, and enrichment for 2 hours (12:20 to 2:20), Monday through Wednesday. On Thursdays, scholars generally attended a field trip in the afternoon.

RQ21. *Of those who completed the program, was their average daily attendance 80 percent or higher?*

To determine the extent to which scholars received the proper dosage of the program, data on enrollment, retention, and attendance were collected at each site and reported to BellXcel headquarters upon completion of both summer programs. These data were then provided in the aggregate to researchers. The indicator of interest related attendance was whether scholars had an average daily attendance of 80% or higher. Table 9 shows 233 and 241 were enrolled in summers 2016 and 2017, respectively. It also shows that the average daily attendance across sites for both years was 80%. In 2016, the only school to achieve average daily attendance of at least 80% was School 4. The other two schools came close to an 80% average attendance rate at 79% (School(1) and 77% (School(3), respectively. In 2017, School 1 and School 4 achieved average daily attendance of at least 80% (82% and 81%, respectively), and School 3 came very close, at 78%. School 4 reported the highest rate of retention (89%), followed closely by School 3 (88%), and School 1 (82%).

Table 9. Enrollment, completion, and attendance data, by site, summers 2016 and 2017

Site	# Enrolled		# Completed		% Overall Average Daily Attendance	
	Summer 2016	Summer 2017	Summer 2016	Summer 2017	Summer 2016	Summer 2017
School 1	63	76	51	62	79%	82%
School 3	85	76	80	67	77%	78%
School 4	85	89	68	79	83%	81%
Total	233	241	199	208	80%	80%

NOTE: "Completed" is defined as a scholar who has participated all the way through the program. A scholar is considered "withdrawn" if he or she signed up for BellXcel, had at least one attendance marked "present," but was removed from the program for any reason (e.g., stopped attending, formally withdrew, or was asked to leave. "Enrolled" is the total of scholars "completed" and "withdrawn." Note that instead of "completed" and "enrolled," BellXcel uses the terms "enrolled" and "served," respectively.
SOURCE: BellXcel Enrollment and Retention data, summer 2016; BellXcel Enrollment and Retention data, summer 2017

Table 10 below breaks down the percentage of scholars in each attendance category, by site. Approximately 64% of scholars in 2016 and 62% of scholars in 2017 across all three sites attended at

least 80% of program days. School 4 reported the highest percentage of students attending 80% of program days for both years, at 69% in 2016 and 65% in 2017.

Table 10. Attendance Rates by site, 2016 and 2017

Site	< 70%		70-79%		80-89%		90 – 94%		95-100%		Total 80% or higher	
	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017
School 1	22	24	16	15	22	23	16	8	25	31	63	62
School 3	28	2	15	15	21	19	9	9	28	31	58	59
School 4	21	18	10	18	25	20	12	18	32	27	69	65
Total	24	22	14	16	23	21	12	12	29	29	64	62

NOTE: Data were provided as rounded percentages; therefore, percentages may sum to more than 100% due to rounding.
SOURCE: BellXcel Enrollment and Retention data, summer 2016; BellXcel Enrollment and Retention data, summer 2017

In focus groups and interviews, staff indicated general satisfaction with scholars' level of attendance in the program, although some sites had mild frustrations. For example, teachers at School 1 reported difficulties with having an inconsistent number of scholars from one day to the next. In particular, teachers at School 1 and School 3 noted that many scholars were on vacation over the 4th of July week so the following week had to be spent on review for the benefit of the absent scholars; teachers suggested that no classes be held over the 4th of July week in future years. All sites indicated that they offered incentives to scholars for perfect attendance, although staff at School 1 noted that such incentives could be problematic if they encouraged scholars to come in sick.

RQ22. *What barriers, if any, prevented scholars from attending the program and/or using transportation at the desired rates? What strategies encouraged attendance and transportation utilization?*

BellXcel provides free bus transportation to scholars for the duration of the summer program to ensure that all who want to attend the program aren't prevented from doing so by lack of access to transport. On the 2017 parent survey, parents of scholars who never rode the bus or only rode it occasionally were asked why those scholars did not ride more frequently. The most frequent responses were that a family member or other adult drove the scholar to the summer program**.** Only two parents indicated that the bus times and/or stop locations were not convenient. No parents indicated that they did not know that a bus was available.[14]

RQ23. *What aspects of the program did scholars prefer over others?*

As seen in Table 11, very high percentages of scholars reported enjoying the field trips, while a slightly smaller percentage reported enjoying the afternoon enrichment classes. Meanwhile, the academic classes were the least liked BellXcel activity, where less than 60% of scholars reported enjoying them across both years.

[14] Two parents selected "other" as the reason their scholar never rode the bus or only rode it occasionally. When asked for an explanation, one parent wrote "didn't make it to the bus" and another indicated driving the scholar to the summer program.

Table 11. Aspects of BellXcel Summer Program scholars enjoyed, by year

	I enjoyed the morning academic classes (n=178)		I enjoyed the afternoon enrichment classes (n=177)		I enjoyed the field trips (n=180)	
	Summer 2016	Summer 2017	Summer 2016	Summer 2017	Summer 2016	Summer 2017
Yes	49%	59%	93%	85%	94%	92%
Maybe	35%	31%	6%	13%	5%	8%
No	15%	10%	1%	2%	1%	0%
Total	100%	100%	100%	100%	100%	100%

Table 12 below breaks down which scholars found math and reading classes interesting. In both years of implementation, students reported enjoying the math classes slightly more, on average, than the ELA classes.

Table 12. Extent to which scholars found math and reading classes interesting, by year

	Math Class		Reading and English Class	
	2016	2017	2016	2017
Not interesting at all	3%	3%	7%	6%
Slightly interesting	16%	4%	15%	4%
Somewhat interesting	11%	18%	24%	22%
Quite interesting	31%	38%	19%	34%
Extremely interesting	39%	37%	35%	35%
Total	100%	100%	100%	100%

The majority of students across the three sites reported that both their math and reading and English skills got slightly better or much better after coming to BellXcel. In general, students enjoyed coming to BellXcel; more than 85% of all students in 2017 enjoyed BellXcel.

Around three-quarters of scholars reported that they got along quite a bit or a tremendous amount with classmates at BellXcel. In contrast, much higher percentages reported that they got along with their teachers quite a bit or a tremendous amount (90%) and that adults at BellXcel cared about them quite a bit or a tremendous amount (80%). Interestingly, a higher percentage of scholars at School 1 reported that adults at BellXcel cared for them a tremendous amount, compared to Schools 3 and 4.

Over 80% of scholars in both years of implementation reported that they felt slightly more confident or much more confident in themselves after the BellXcel Summer Program. Overall, most scholars indicated that teachers were encouraging and gave them the feedback they needed to help them learn. In general, scholars wanted to see more, or improved, field trips and less work/more play time. Many students from both years indicated that they did not want to change anything about the BellXcel program.

RQ24. *What suggestions do scholars have for improvement of the various program components as implemented?*

On the scholar survey in summer 2017, scholars were asked to write in their suggestions to make the overall BellXcel experience better. The most common response was around field trips. Fourteen respondents indicated that they would like the summer program to have more field trips, with a couple of scholars indicating that they would also like the field trips to be improved.

Thirteen scholars mentioned a need for less work and more play, with many noting a desire for more "activities" and less "classroom time." Seven scholars mentioned a desire for more and/or improved enrichment activities. Three scholars indicated they wanted more reading instruction, while two requested more or better math instruction. Three scholars requested better scholar dollar prizes. Two scholars wished for more days in the summer program. A total of 21 scholars indicated not wanting to change anything about the summer program. Two scholars noted their appreciation for the opportunity to meet new friends at BellXcel.

E. Parent Engagement Activities and Outputs

RQ25. *Did at least 70 percent of parents attend parent engagement events?*

Parent engagement events are an important part of the BellXcel Summer Program. In Greenville, each site conducted a parent orientation, an open house, a family literacy workshop, and a closing ceremony. The logic model specifies a goal of 70% of parents attending parent engagement events.

No site reached 70% parent participation at any one event in either 2016 or 2017 (Table 13). In 2016, parent participation rates varied from as low as 6 percent to as high as 43 percent. In 2017, parent participation rates varied from as low as 7% to as high as 51%. Unfortunately, it is impossible to surmise from the attendance data if the same parents attended multiple events or if different parents attended different events.

Table 13. Number and percent of scholar families attending parent engagement activities, by site

	Parent Orientation		Open House		Family Literacy Workshop		Closing Ceremony	
Site	**Summer 2016**	**Summer 2017**	**Summer 2016**	**Summer 2017**	**Summer 2016**	**Summer 2017**	**Summer 2016**	**Summer 2017**
School 1	12 (24%)	25 (32%)	7 (14%)	10 (16%)	4 (8%)	N/A	13 (25%)	25 (40%)
School 3	15 (19%)	30 (14%)	14 (18%)	20 (30%	5 (6%)	5 (7%)	15 (19%)	25 (37%)
School 4	19 (28%)	30 (34%)	17 (25%)	25 (32%)	7 (10%)	N/A	29 (40%)	40 (51%)

NOTE: Percent attendance was calculated by dividing the number of scholar families by the total number of scholars enrolled at each site. Therefore, percent attendance should be viewed as an approximation only.
SOURCE: BellXcel Summer 2016 Tracking Sheet for Family Events; BellXcel Summer 2017 Tracking Sheet for Family Events

When asked during focus groups about parent engagement in BellXcel, staff uniformly reported it was relatively strong and that BellXcel had many avenues for encouraging parents to connect with the program. In addition to the events noted in Table X, each site sent home a weekly newsletter, made robo-calls for all events, and conducted daily attendance phone calls whenever a scholar was absent.

A very small number of parents responded to the questions about why they did not attend the parent orientation or family literacy workshop. Of those who responded, the most frequently cited reason for not attending the parent orientation was a scheduling conflict, while for the family literacy event[15] it was lack of awareness (Table 14).

In 2016, staff at School 1 indicated that they believed that transportation was a large barrier to parent attendance. They indicated that many parents do not have access to a car, and because the school accepts scholars from 15 feeder elementary schools, families tend to be spread out. At School 4, by contrast, a staff member indicated that transportation may be an issue for some, but not for most families.

In 2017, five parents indicated that they had trouble finding transportation or childcare for the family literacy workshop, and two parents provided this reason for not attending the parent orientation. One parent notably indicated, in "other" responses, not being able to attend either event due to lack of wheelchair accessibility.

Table 14. Reasons parents gave for not attending a parent engagement event

Reason for not attending (check all that apply)	**Parent orientation – Number of parents**		**Family literacy workshop – Number of parents**	
	Summer 2016	Summer 2017	Summer 2016	Summer 2017
Another family member attended	0	0	0	2
I did not think it would be useful	0	0	0	0
Trouble finding transportation or childcare	0	2	0	5
Other	4	5	6	5
I was not aware of the event	5	9	7	13
Schedule conflict	7	11	9	8

SOURCE: BellXcel Parent Survey, summer 2016; BellXcel Parent Survey, summer 2017

RQ26. *How effective were the parent engagement activities?*

Parent survey respondents who attended the parent engagement events generally had positive feedback. For both years of implementation, at least 70 percent of parents agreed that each event helped them to understand their scholar's experience in the BellXcel program; that they were well planned with a clear agenda; and that they provided the opportunity to connect with BellXcel teachers and staff.

[15] The family literacy event was referred to as the family week event in the surveys.

RQ27. *How might parent engagement events be improved?*

The program manager at School 4 mentioned doing a phone blast to encourage participation; it is unclear if School 1 and School 3 used this technique as well. The program manager at School 4 also recommended more of an emphasis on parent involvement in the future, suggesting doing more regular events for parents, perhaps on a set schedule. Other staff emphasized that night events were better for parents due to work schedules.

The only suggestion for improvement that parents provided on the survey was that they be done on a different schedule so more parents can attend, and that a call or letter inform parents about the events.

IV. Statistical Analysis of Impacts

A. Unit of Assignment and Analysis

The unit of assignment for this study was the individual student. Individual students were identified and selected for participation in the BellXcel Summer Program intervention through targeted recruitment efforts at OnTrack Greenville feeder elementary schools. Researchers adjusted their analysis on the effects of the BellXcel Summer Program to match the unit of assignment by using propensity matching at the student-level. For each student participating in the BellXcel Summer Program, up to five "match" students were selected from each comparison group based on several covariates and background conditions. Thus, comparison and treatment groups were comprised of individual matches, and the unit of analysis reached the student-level.

B. Analysis Approach

The analysis described here followed a Treatment on Treated (TOT) framework, as BellXcel Summer Program defines the treatment. It compared those who completed the BellXcel Summer Program to students in the within-school, within-district, and state comparison groups. This study examined whether students who completed the BellXcel Summer Program exhibited improved course performance compared students who did not participate in the program. Evaluators also examined the relationship between the BellXcel Summer Program and student attendance and behavior. Differences between the treatment and the within-school comparison group were estimated separately from the treatment and the external comparison groups.

C. Formation of Matched Groups

In order for matching techniques to approximate a random experiment, important assumptions have to be met. The first was strongly ignorable treatment assignment. This means that conditional on observed covariates (*X*) the treatment (*W*) was independent of the outcomes (Y_0, Y_1), or $(Y_0, Y_1) \perp W|X$ (Rosenbaum and Rubin 1983; Guo and Fraser 2014, 209). For this assumption to hold, the selection process had to be derived from covariates used in the model. Previous research indicates that results from matching designs only reflect randomized control trial results when the covariates in the propensity score model accurately predict treatment assignment (Bifulco, 2012; Cook et al., 2008; Fortson et al., 2012). Matching methods work best when pretreatment outcome measures are used in estimating the propensity score. For this analysis, the propensity score (*P(X)*) was equivalent to:

$$P(X) = \Pr(T_i = 1|\boldsymbol{X}_i),$$

where T_i = 1 if the student, *i,* completes the BellXcel Summer Program and $\boldsymbol{X}_i$ is a vector of covariates that predict participation in the BellXcel Summer Program. Evaluators selected the covariates that best predicted treatment assignment and imbalance between treatment and control groups. Given the importance of pretreatment outcome measures, attendance (percentage of days attended), behavior (in school and out of school suspensions), and academic performance (math, ELA, social studies, and

science test scores) from the previous academic year were used.[16] In addition to these factors, the propensity score model included students' race, gender, English-language learner status, disability status, low-income indicator, and grade. Previous research indicates that these variables are related to student attainment and student achievement (e.g., Goldschmidt & Wang, 1999; Laird, Kienzi, DeBell, & Chapman, 2007; Reardon & Robinson, 2007; Reschly & Christenson, 2006; Rumberger & Lim, 2008; Stetser & Stillwell, 2014). Therefore, these factors were included in the propensity model, as they could have been predictive of the likelihood of students being identified by the EWRS, the likelihood of completing the BellXcel Summer Program, and the outcomes of interest.

Researchers estimated a logistic regression using these covariates to produce a predicted probability of receiving treatment for each student separately for the within-school comparison group and the external comparison groups. The model used by evaluators was:

$$\Pr(T_i = 1|X) = \exp(\beta \boldsymbol{X}_i)/(1 + \exp(\beta \boldsymbol{X}_i)),$$

where $\boldsymbol{X}_i$ is a vector of covariates discussed above.

After creating propensity scores, treatment students were matched to comparison students. Like the estimation of the propensity scores, the matching between treatment and within-school comparison students and between treatment and external comparison students occurred independently. Evaluators used nearest neighbor matching, which is a form of greedy matching. Treatment students were matched to the comparison student with the closest absolute propensity score, as long as the distance between the propensity scores fell within a caliper of $0.25\sigma_p$, where σ_p is the standard deviation of the propensity scores (Guo and Fraser 2014, 147). This was done to ensure good matches between treatment and comparison students.

Ideally, each BellXcel treatment student was matched to five comparison students to boost sample size. Matching was done with replacement. However, there were instances in which treatment students could not be matched. First, if a student had missing data for any of the variables included in the first stage regression, the student was excluded from the analysis. Second, BellXcel scholars who did not have any potential matches within the caliper described above were not included in the analyses presented in this report, as researchers were not able to identify suitable matches.

Another assumption of the matching methods is that there is proper overlap in the propensity scores between the treatment and control group (Rosenbaum & Rubin, 1983; Stuart & Rubin, 2008). In order to ensure that this assumption is met, evaluators performed bivariate tests, such as *t*-tests or chi-square tests, before and after matching. If these tests revealed a significant level of imbalance or a lack of overlap, then evaluators, following Rosenbaum and Rubin (1984, 1985), re-estimated the propensity model using higher-order polynomial terms and interactions between the covariates. When considering the balance of the matches, researchers examined the statistical significance of the bivariate differences in the post-matching covariates between the treatment and control groups, the standardized mean differences between the two groups, and the variance ratios. The goal was to have no statistically significant differences, standardized mean differences below 0.1, and variance ratios near 1.0 (Steiner & Cook, 2013).

[16] When available, test scores from both MAP and SC READY exams were used. Collinearity between predictors is generally not a threat when estimating the propensity score (Stuart, 2010).

The pre-matching differences between BellXcel students and the comparison students are presented below, followed by data on the effectiveness of the matching procedure.

1. Characteristics of BellXcel Summer Program Students: Pre-Matching

This study examines students who participated in the BellXcel Summer Program in summer 2016 and summer 2017. Table 15, below, presents the demographic characteristics of four groups of students: (1) the 104 students who attended the 2016 BellXcel Summer Program, (2) the general sixth grade student population of treatment schools, (3) the general sixth grade student population of district schools, and (4) the general sixth grade student population of state schools. These latter three groups include all-non BellXcel scholars in the comparison group populations, allowing for a comparison of demographic characteristics before the matching process.

Tables 15 and 16, below, show the demographic characteristics of BellXcel scholars and the student populations of comparison schools in academic years 2016-17 and 2017-18. As shown in these tables, BellXcel scholars varied significantly in demographics from students at treatment, district, and state schools.

Table 15. Pre-Matching Demographic Characteristics AY 2016-17

	BellXcel Scholars (n = 104)	Student Population: Treatment Schools (n = 535)	Student Population: District Schools (n = 1,162)	Student Population State Schools (n = 16,547)
Black	41.7%	**30.9%*** (-0.23)	**27.6%**** (-0.30)	**49.2%†** (0.05)
Hispanic	45.6%	**32.9%*** (-0.26)	**16.9%***** (-0.65)	**9.7%***** (0.03)
White	10.7%	**27.4%***** (0.44)	**47.9%***** (0.90)	**41.2%***** (0.05)
Other Race	1.9%	**8.8%*** (0.31)	**7.6%*** (0.27)	4.3% (0.13)
Free or Reduced Meals	95.1%	93.3% (-0.08)	**56.8%***** (-1.00)	**75.7%**** (0.04)
Female	53.4%	46.0% (-0.15)	**44.9%**† (-0.17)	48.4% (0.05)
Special Education	26.2%	21.5% (-0.11)	**12.3%***** (-0.36)	**15.2%*** (0.04)
ESL	38.8%	**25.9%**** (-0.28)	**13.3%***** (-0.61)	**6.1%***** (0.02)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Note: Standardized mean differences are reported in parentheses. Significance tests are in comparison to the BellXcel group.

As shown in Table 14, 42% of BellXcel scholars were Black, 46% were Hispanic, 11% were White, and 29% were some other race in academic year 2016-17. BellXcel scholars were more likely to be Hispanic than sixth grade students in the treatment schools ($p < 0.05$), district schools ($p < 0.001$), and state schools ($p < 0.001$). In addition, BellXcel scholars were less likely to be White than students in treatment schools ($p < 0.001$), district schools ($p < 0.001$) and state schools ($p < 0.001$). Nearly all BellXcel scholars

(95%) and the treatment school sixth grade students (93%) were low-income students, as measured by free and reduced meals, while only 57% of sixth grade students in the district schools ($p < 0.001$) and 76% of sixth grade students in state schools ($p < 0.01$) were low-income. More BellXcel scholars were designated as special education than the general student population of district schools, ($p < 0.001$) and state schools ($p < 0.05$). BellXcel scholars were more likely to be designated ESL than sixth grade students in treatment schools ($p < 0.01$), district schools ($p < 0.001$) and state schools ($p < 0.001$). Many of these pre-matching demographic differences were similar in academic year 2017-18. See Table 16.

Table 16. Pre-Matching Demographic Characteristics AY 2017-18

	BellXcel Scholars (n = 106)	Student Population: Treatment Schools (n = 505)	Student Population: District Schools (n = 1,184)
Black	29.2%	28.5% (0.02)	29.6% (0.01)
Hispanic	55.7%	**45.1%*** (0.21)	**14.4%***** (0.96)
White	12.3%	**20.5%†** (0.22)	**48.6%***** (0.86)
Other Race	2.8%	5.9% (0.15)	**7.4%†** (0.21)
Free and Reduced Meals	86.8%	84.9% (0.06)	**59.5%***** (0.65)
Female	50.9%	44.6% (0.13)	48.7% (0.05)
Special Education	19.8%	**12.3%*** (0.21)	**7.6%***** (0.36)
ESL	31.1%	28.0% (0.07)	**7.7%***** (0.62)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Note: Standardized mean differences are reported in parentheses. Significance tests are in comparison to the BellXcel group.

In addition to the pre-match demographic characteristics shown above, it is worthwhile to examine how BellXcel scholars compared to the school populations in school achievement outcomes from academic years 2015-16 and 2016-17, the year prior to participation in the BellXcel Summer Program. These pre-treatment outcomes were used in the propensity score matching process and as covariates in outcome regression analyses. A selection of prior year outcomes appears in Table 17 below.

Scholars who participated in the summer 2016 BellXcel Summer Program had very similar behavior outcomes as the overall sixth-grade student population of treatment schools. However, BellXcel scholars had higher rates of average daily attendance ($p < 0.001$) and lower rates of chronic absenteeism ($p < 0.10$) when compared to sixth-grade students attending treatment schools. BellXcel scholars also had lower prior year ELA ($p < 0.05$) and math scores ($p < 0.10$) on the SC READY assessment than sixth-grade students at treatment schools.

When compared to all sixth grade students attending district schools, BellXcel scholars were more likely to have received any behavioral referral ($p < 0.05$), were more likely to have received out-of-school

suspension ($p < 0.10$), and had higher rates of average daily attendance ($p < 0.05$) in the 2015-16 academic year. BellXcel scholars also had lower prior year ELA ($p < 0.001$) and math scores ($p < 0.001$) on the SC READY assessment than sixth-grade students at district schools. When compared to students at state schools, BellXcel scholars differed significantly on every measure except for the proportion of students receiving at least one behavioral referral in the 2015-16 academic year. Many of the significant differences in pre-matching prior year outcomes that were present in summer 2016 also were significant in summer 2017, as shown in Table 18. These differences provide additional support for the importance of using propensity score matching in this study.

Table 17. Pre-Matching Prior Year Outcomes for Summer 2016

	Summer 2016 BellXcel Scholars (n = 103)	Student Population: Treatment Schools (n = 511)	Student Population: District Schools (n = 1,117)	Student Population: State Schools (n = 16,526)
Any Behavioral Referral	27.9%	26.2% (0.05)	**19.1%*** (0.04)	29.9% (0.08)
Any ISS	1.9%	5.2% (0.02)	1.2% (0.01)	**8.1%**** (0.42)
Any OSS	8.7%	10.5% (0.03)	**4.9%**†** (0.02)	**11.7%***** (0.51)
SC READY – ELA	1533	**1538*** (-0.25)	**1546***** (-0.67)	**1540**** (-0.33)
SC READY - Math	1539	**1543**†** (-0.21)	**1551***** (-0.63)	**1543*** (-0.22)
Average Daily Attendance	97.5%	**95.7%***** (0.48)	**96.9%*** (0.30)	**95.9%***** (0.43)
Chronic Absenteeism	3.0%	**8.0%**† (0.03)	3.0% (0.02)	**8.4%*** (0.24)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Note: Standardized mean differences are reported in parentheses. Significance tests are in comparison to the BellXcel group.

Table 18. Pre-Matching Prior Year Outcomes for Summer 2017

	Summer 2017 BellXcel Scholars (n = 106)	Student Population: Treatment Schools (n = 505)	Student Population: District Schools (n = 1,184)
Any Behavioral Referral	23.9%	26.7% (0.07)	19.0% (0.12)
Any ISS	3.7%	6.7% (0.14)	1.9% (0.11)
Any OSS	9.2%	8.7% (0.02)	**5.2%**† (0.16)
SC READY – ELA	1536	1538 (-0.12)	**1545***** (-0.51)
SC READY – Math	1539	**1544*** (0.24)	**1550***** (-0.53)
Average Daily Attendance	97.3%	**95.7%***** (-0.52)	**96.2%**** (-0.36)
Chronic Absenteeism	1.0%	**7.6%*** (0.33)	**5.4%*** (0.25)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Note: Standardized mean differences are reported in parentheses. Significance tests are in comparison to the BellXcel group.

2. Effectiveness of the Matching Procedure

To assess the impact of BellXcel Summer Program participation on student course performance and student social-emotional outcomes, researchers created seven matched comparison groups. For the 2016-17 analyses, three matches were performed, one for each comparison group (treatment school, district, and state), using the administrative and survey data. In academic year 2017-18, the research team decided to do separate matches for the administrative and survey data, as a fair number of students did not complete the survey. Two separate matches were completed for both the treatment school and district comparison with the 2017-18 data.

Researchers re-estimated the propensity model until balance was achieved between the treatment and comparison groups. Balance was determined by examining the statistical significance of the bivariate differences, the magnitude of the standardized mean differences, the variance ratios, and the distribution of the propensity scores between the two groups. To provide an example of the success of the matching process, the results for the 2017-18 administration data match between BellXcel students and the within-school treatment school comparison students are presented below.

Figure 1 demonstrates that the matching process produced two similar groups in terms of standardized differences. For all the 16 variables in the analysis, none have a standardized mean difference larger than 0.1. Figure 2 provides evidence that the two groups are also similar in terms of the variance of the variables. Further, there is substantial overlap between the BellXcel students and the treatment school matches in terms of their propensities to participate in BellXcel, as is evident in Figure 3. Given these results, researchers are confident that the two groups are similar on these observable characteristics.

Figure 1. 2017-18 Treatment School Comparison Matches (Administrative Data): Standardized Differences

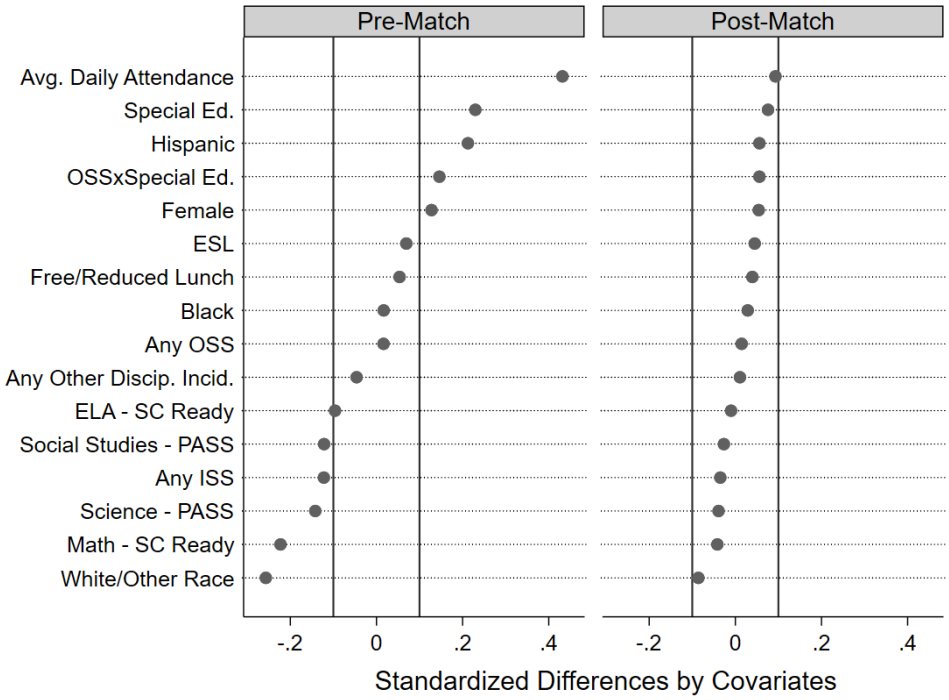


Figure 2. 2017-18 Treatment School Comparison Matches (Administrative Data): Variances

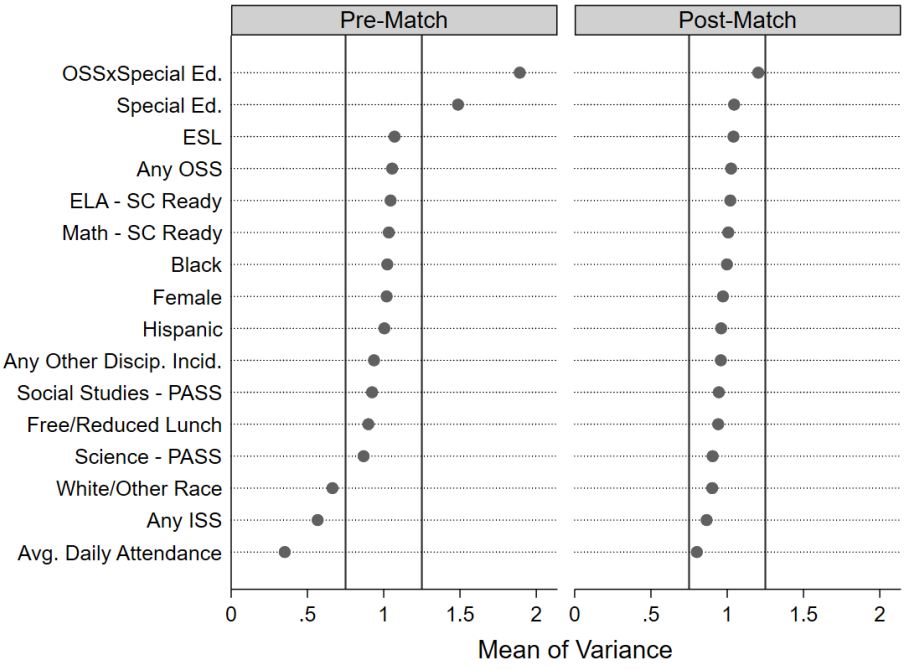
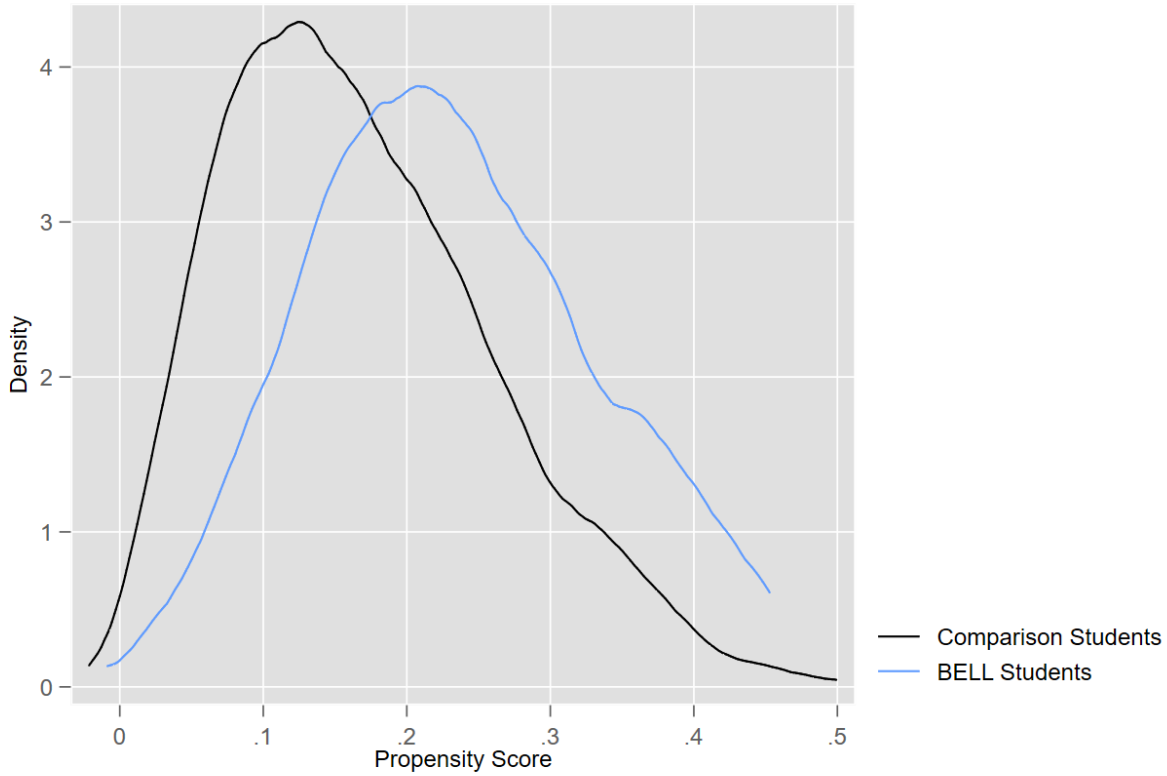


Figure 3. 2017-18 Treatment School Comparison Matches (Administrative Data): Overlap



A similar method was used for the six other matching procedures. In the instances in which ideal matches were not possible, the differences were often marginal. For the results of the other matches, see Appendix C.

3. Characteristics of BellXcel Summer Program Students: Post-Matching

It was necessary to create distinct comparison groups due to (1) the three different comparison school populations; and (2) the two different sources of data, administrative and survey. The following series of tables presents the demographic characteristics of BellXcel Summer Program students and their matched counterparts for each school (treatment, district, or state) and each type of outcome data.

As noted below in Table 19, no significant differences were present between treatment students and matched comparison students attending treatment and district schools on demographic factors for the academic year 2016-17 matches. The lack of significant differences here indicates that the propensity score matching process resulted in balanced samples. More sophisticated analyses of standardized differences and variance ratios between the treatment and comparison students confirmed that the samples were balanced when examining demographic factors, as well as the baseline attendance, behavior, and achievement outcomes used in the propensity score model. For more information on the effectiveness of the matches for other data sources and academic years, see Appendix C.

Table 19. BellXcel Summer Program Participants vs. Student Matches 2016–17 (Administrative and Survey Data Match)

	BellXcel Participants (n = 100)	Student Matches: Treatment Schools (n = 500)	BellXcel Participants (n = 89)	Student Matches: District Schools (n = 445)	BellXcel Participants (n = 98)	Student Matches: State Schools (n = 465)
Black	41.0%	45.2% (-0.06)	46.1%	46.7% (-0.01)	41.8%	40.6% (0.02)
Hispanic	46.0%	45.4% (0.03)	39.3%	39.8% (0.01)	44.9%	44.5% (0.01)
White	11.0%	9.0% (0.07)	12.4%	11.0% (0.04)	13.3%	14.9% (-0.05)
Other Race	2.0%	0.4% (-0.03)	2.3%	2.5% (-0.01)		
Free and Reduced Meals	96.0%	96.2% (-0.05)	94.4%	95.1% (-0.03)	89.8%	90.8% (-0.03)
Female	53.0%	52.4% (0.01)	52.8%	50.6% (0.04)	52.0%	52.0% (0.00)
Special Education	26.0%	23.4% (0.01)	25.8%	25.6% (0.01)	21.4%	24.7% (-0.07)
ESL	39.0%	38.4% (0.03)	31.5%	29.0% (0.05)	34.7%	35.5% (-0.02)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Note: Standardized mean differences are reported in parentheses. Significance tests are in comparison to the BellXcel group. Frequency weights were used to account for matching with replacement.

Researchers were not able to identify suitable matches for some BellXcel participants due to the parameters of the propensity score matching process. This means that not all BellXcel scholars were included in the analyses presented in this report. Tables 20 and 21, below, describe how many BellXcel scholars were matched for each comparison group and each set of outcome data. The total number of students listed in the table includes all BellXcel students who met the inclusion criteria and have full matching data.

When looking at the 101 BellXcel Scholars who were included in the treatment group in academic year 2016-17, 99% of participants were matched to comparison students at treatment schools and district schools. Similarly, when looking at the scholars who were included in the treatment group and completed the OnTrack Greenville student survey in 2017-18, 93% of BellXcel participants were matched to comparison students at treatment schools and 98% were matched to district schools students. Further, matching proportions were relatively consistent from academic year 2016-17 to 2017-18.

It is important to note that the estimated effects of the program only pertain to those students included in the analyses. It is possible that participating in the BellXcel Summer Program had different effects on those students who were not matched. However, the research team was able to match the vast majority of BellXcel scholars to lessen this problem.

Table 20. Number and Percent of BellXcel Scholars Matched (Administrative Data)

	Academic Year 2016-17						Academic Year 2017-18			
	Treatment Schools		District Schools		State Schools		Treatment Schools		District Schools	
Matched	100	99%	100	99%	98	94%	100	94%	101	95%
Not Matched	1	1%	1	1%	3	6%	6	6%	5	5%
Total	101	100%	101	100%	101	100%	106	100%	106	100%

Table 21. Number and Percent of BellXcel Scholars Matched in (Survey Data)

	Academic Year 2017-18			
	Treatment Schools		District Schools	
Matched	74	93%	78	98%
Not Matched	6	7%	2	2%
Total	80	100%	80	100%

It is important to note that the estimated effects of the program only pertain to those students included in the analyses. It is possible that participating in the BellXcel program had different effects on those students who were not matched. However, the research team was able to match the vast majority of BellXcel students to lessen this problem.

To examine how representative the matched set of BellXcel scholars is, researchers compared the student demographics between the matched BellXcel students and those BellXcel students who could not be matched because they did not have full matching data (refer to Table 5) or did not have a proper match within the given caliper. The results for the 2017-18 administrative data match with the treatment school comparison group appears in Table 22 below. Appendix C provides information on the demographic differences between those BellXcel students who were matched and those who were not.

Table 22. Which BellXcel Students Were Matched? Post-Match Demographics, 2017-18 Treatment School Administrative Data Match

	BellXcel Not Matched - Mean	BellXcel Matched - Mean	Difference	S.E. of Diff.
Black	0.33	0.29	0.04	0.19
Hispanic	0.67	0.55	0.12	0.21
White	0.00	0.13	-0.13	0.14
Other Race	0.00	0.03	-0.03	0.07
Free/Reduced Meals	0.83	0.87	-0.04	0.14
Female	0.50	0.51	-0.01	0.21
Special Ed.	0.83	0.16	**0.67*****	0.16
ESL	0.33	0.31	0.02	0.20

Significance determined by chi-square tests. + 0.10 * 0.05 ** 0.01 *** 0.001

D. Impact Study Results

There are many competing matching procedures. Researchers for this study employed greedy matching. While other approaches, such as optimal matching, have their strengths, greedy matching allowed researchers to perform multivariate analyses on the samples after matching. This feature is one of the reasons why greedy matching is so popular across many disciplines (Guo & Fraser, 2014, p. 148). When propensity scores are used with a regression adjustment, the estimates are "doubly robust," which helps with robustness against misspecification in the propensity score model or the regression model (Imbens & Wooldridge, 2009).

As specified above, researchers ensured that there were no significant differences between the treatment and comparison groups on pre-treatment covariates. For each outcome, the basic model to estimate the impact effects of the BellXcel Summer Program was as follows:

$$Y_{t,i} = \beta_0 + \beta_1 T_i + \beta_2 Y_{Baseline,i} + \beta_3 R_i + \beta_4 F_i + \beta_5 I_i + \beta_6 D_i + \beta_7 E_i + e_i$$

In this equation, $Y_{t,i}$ is the outcome for student *i* at time *t*. β_1 represents the impact of completing the BellXcel Summer Program (T =(1)). A statistically significant estimate of β_1 indicates that participating in the BellXcel Summer Program is related to the outcome of interest. When estimating impact effects using a quasi-experimental design, inclusion of a pretreatment outcome measure decreases selection bias and increases precision (e.g, Bifulco, 2012). β_2 is the impact of the pretreatment, or baseline outcome.[17] For example, the regression model predicting fall 2016 math MAP assessment scores will include the spring 2016 MAP math score as a covariate. R_i represents a set of dummy variables for race (Black, Hispanic, and other with white the omitted category). β_4 represents the difference between female and male students, and I_i is a dummy variable indicating if a student is eligible for free or reduced-priced lunch. Dummy indicators for disability status (D_i) and English as a second language status (E_i) were also included in the model.

The student-level random error is denoted as e_i in the above model. One assumption of a traditional OLS model is that the residuals are uncorrelated with the covariates. Since this assumption of homoscedasticity does not necessarily hold, the models were estimated with robust standard errors. Bootstrap methods were used to estimate the standard errors. This approach relies on sampling from the analysis sample and replicating the analysis. This study used 500 replications to produce the standard errors of the BellXcel coefficients.

The Stata software program was used to perform the matching procedure and outcome analyses (StataCorp, 2017). The above model was used for interval dependent variables, such as test scores, while a logistic regression was estimated for dichotomous dependent variables, like whether or not a student was chronically absent during a given semester or school year. The student self-reported frequency of summer activities measures are ordinal; therefore, ordered logistic regression was used for these outcomes.

While the main impact analyses compared BellXcel scholars to matched comparison students, researchers also estimated regressions using the population of comparison students. These unmatched

[17] Whenever possible, the same measure was used for both the outcome and the pretreatment control variable. When that was not possible, a pretreatment or baseline measure from the same domain as the outcome variable was used. No pretreatment measures of the outcome were controlled for the analyses of the survey data.

regression results allow one to compare the overall student populations in each comparison group to students who participated in the BellXcel program in summer 2016 or 2017. These unmatched analyses provide context when assessing the impact of the matching procedure on the overall conclusions of the study. The results of these unmatched regression analyses are provided in Appendix J.

1. Estimates of Effect Size

Evaluators of education programs should consider the substantive impact of interventions in addition to their statistical significance. For the impact analyses presented below, researchers converted the multivariate regression coefficients into covariate-adjusted, standardized effect sizes (Cohen's d). This was a straightforward process for the OLS regression coefficients (Lipsey & Wilson, 2000). To convert the binary and ordered logistic regression results to Cohen's d estimates, researchers first transformed the coefficients to odds ratios. Following Borenstein, Hedges, Higgins, and Rothstein (2009), these odds ratios were then converted into estimates of Cohen's d (p. 47). A benefit of Cohen's d is that it allows for comparisons of substantive impacts across outcomes and studies. However, there is no agreed upon definition of a "meaningful" effect size. Cohen (1988) suggested that an effect size of 0.2 should be considered small, 0.5 moderate, and 0.8 large. However, effect sizes of this magnitude are quite rare in education evaluations. Perhaps, a more appropriate threshold is the What Works Clearinghouse's (2017) statement that an effect size of 0.25 or greater should be considered "substantively important."

2. Confirmatory Impact Results

RQ1. *Did students who participated in the BellXcel Summer Program demonstrate improved math and English/language arts course performance when compared to matched comparison students?*

In this study, the effect of BellXcel participation on MAP assessment reading and math test scores were confirmatory analyses. Table 23 presents the differences in ELA and math achievement between BellXcel and matched comparison students at treatment and district schools. MAP assessment scores were not available for state comparison students in academic year 2016-17 and were not available at all in academic year 2017-18. Two different sets of data are presented in the following tables. The "Matched Regression" columns show the BellXcel regression coefficients from the post-match multivariate regressions and the bootstrapped standard errors. The regression coefficients from the OLS regressions are directly interpretable. For analyses of dichotomous and ordinal dependent variables, binary and ordered logistic regression was used, and the "Matched Regression" column presents the log odds and associated bootstrapped standard errors. The "Effect Size" column displays the Cohen's d estimate associated with the BellXcel regression coefficient. Positive values in the table indicate that the BellXcel scholars exhibited greater growth for an outcome than the comparison group. Statistically significant differences between the groups are denoted by asterisks in the tables.

As shown in Table 23, there were no significant differences in ELA or math course performance as measured by the MAP assessment between BellXcel scholars and matched comparison students at treatment and district schools in academic year 2016-17.

Table 23. Confirmatory Impact Results for Course Performance AY 2016-17

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools		BellXcel vs. Comparison Students: State Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size	Matched Regression	Effect Size
MAP RIT - Math Fall 2016	0.14 (0.78)	0.01	-0.47 (0.91)	-0.04	---	---
MAP RIT - Reading Fall 2016	-1.00 (1.23)	-0.06	-1.25 (1.25)	-0.08	---	---

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; BellXcel vs. Treatment N = 600; BellXcel vs. District N = 534

Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measures. Frequency weights were used to account for matching with replacement.

3. Exploratory Impact Results

RQ2. *Did students who participated in the BellXcel Summer Program have higher school attendance rates than matched comparison students?*

Researchers examined differences in attendance between BellXcel scholars and matched comparison students. The matched analyses adjusted for student attendance in the prior academic year before participation in the BellXcel Summer Program, in addition to student demographic data. For the analyses examining average daily attendance, positive values indicate that BellXcel scholars exhibited a more desirable outcome than matched comparison students. For the analyses examining chronic absenteeism, negative values in the table indicate that BellXcel scholars exhibited a more desirable outcome (e.g., less likely to be chronically absent) than matched comparison students. The exploratory impact results for attendance are presented in Tables 24 and 25.

Table 24. Exploratory Impact Results for Attendance AY 2016-17

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools		BellXcel vs. Comparison Students: State Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size	Matched Regression	Effect Size
Average Daily Attendance	0.13 (0.35)	0.03	0.52 (0.40)	0.11	0.22 (0.34)	0.05
Chronically Absent	-0.71 (0.71)	-0.39	-0.55 (0.56)	-0.30	-0.32 (0.49)	-0.17

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ BellXcel vs. Treatment School N= 588; BellXcel vs. District N= 528; BellXcel vs. State N = 588

Note: Table presents the regression coefficients of the BellXcel variable and and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measures. Frequency weights were used to account for matching with replacement.

In academic year 2016-17, there were no significant differences in average daily attendance or chronic absenteeism between summer 2016 BellXcel scholars and their patched counterparts at treatment, district, or state schools. See Table 24. In academic year 2017-18, scholars who participated in the BellXcel Summer Program in summer of 2017 had a significantly higher average daily attendance in academic year 2017-18 than matched comparison students attending treatment schools ($p < 0.01$) and district schools ($p < 0.01$), as shown in Table 25.

Table 25. Exploratory Impact Results for Attendance AY 2017-18

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size
Average Daily Attendance	**0.98**** (0.33)	0.27	**0.94**** (0.34)	0.28
Chronically Absent	-0.43 (0.70)	-0.24	-0.39 (0.78)	-0.22

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; BellXcel vs. Treatment N = 600; BellXcel vs. District N = 606
Note: Table presents the regression coefficient of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure, when possible.

RQ3. *Did students who participated in the BellXcel Summer Program have fewer behavioral incidences than matched comparison students?*

Researchers examined differences in behavior between BellXcel scholars and matched comparison students. The matched analyses adjusted for student behavior in the year before participation in the BellXcel Summer Program, in addition to student demographic data. For the analyses of in-school suspension (ISS), out-of-school suspension (OSS), and behavioral referrals, negative values in the table indicate that BellXcel scholars exhibited a more desirable outcome (e.g., less likely to have any in-school suspension) than matched comparison students.

Table 26 presents the exploratory impact results for behavior for academic year 2016-17. There were no significant differences between BellXcel scholars and matched comparison students attending treatment schools. There were, however, several significant differences between BellXcel scholars and matched comparison students attending district schools on behavioral outcomes in the 2016-17 academic year. When compared to matched students at district schools, BellXcel scholars received an average of 0.64 fewer behavioral referrals ($p < 0.05$), 0.96 fewer days of out-of-school suspension ($p < 0.01$), and 5.24 fewer hours of in-school suspension ($p < 0.01$) than matched students at district schools. Also, BellXcel scholars were less likely to receive any in-school suspension than matched comparison students at district schools during the 2016-17 academic year ($p < 0.05$). These results suggest that BellXcel scholars had better behavioral outcomes than matched comparison students at district schools during the 2016-17 academic year. In general, BellXcel scholars and matched comparison students at state schools had similar behavior outcomes. The only exception was that BellXcel scholars did receive 0.26 fewer in-school suspensions than matched comparison students at state schools ($p < 0.10$).

Table 26. Exploratory Impact Results for Behavior AY 2016-17

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools		BellXcel vs. Comparison Students: State Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size	Matched Regression	Effect Size
Any Behavioral Referral	0.18 (0.30)	0.10	0.26 (0.31)	0.14	0.17 (0.27)	0.10
# Behavioral Referrals	-0.47 (0.30)	-0.15	**-0.64*** (0.27)	-0.19	-0.31 (0.28)	-0.08
Any ISS	0.05 (0.41)	0.03	**-0.92*** (0.36)	-0.51	-0.35 (0.35)	-0.19
# Hours ISS	0.02 (0.72)	0.00	**-5.24**** (1.16)	-0.41	**-0.26†** (0.14)	-0.16
Any OSS	0.45 (0.38)	0.25	-0.24 (0.34)	-0.13	0.08 (0.32)	0.04
# Days OSS	-0.22 (0.19)	-0.09	**-0.96**** (0.29)	-0.26	-0.08 (0.10)	-0.06

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ BellXcel vs. Treatment School N= 588; BellXcel vs. District N= 522; BellXcel vs. State N = 588

Note: Table presents the regression coefficients of the BellXcel variable and and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measures. Frequency weights were used to account for matching with replacement.

Table 27. Exploratory Impact Results for Behavior AY 2017-18

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size
Any Behavioral Referral	-0.12 (0.30)	-0.06	0.29 (0.31)	0.16
# Behavioral Referrals	-0.64 (0.40)	-0.14	-0.02 (0.31)	-0.01
Any ISS	-0.20 (0.36)	-0.11	-0.21 (0.36)	-0.12
# Hours ISS	-0.91 (0.78)	-0.13	**-2.91**** (1.16)	-0.23
Any OSS	**-0.82*** (0.38)	-0.45	-0.27 (0.41)	-0.15
# Days OSS	**-0.73**** (0.28)	-0.23	**-0.59⁺** (0.33)	-0.18

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; BellXcel vs. Treatment N = 600; BellXcel vs. District N = 606

Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measures. Frequency weights were used to account for matching with replacement.

In academic year 2017-18, there were several significant differences between BellXcel scholars and matched comparison students on behavioral outcomes, shown above in Table 27. When compared to matched students at treatment schools, BellXcel scholars were less likely to have received any out-of-school suspension ($p < 0.05$) and they received an average of 0.73 fewer days of out-of-school suspension than matched comparison students at treatment schools ($p < 0.01$). When compared to the matched comparison students attending district schools, BellXcel scholars received an average of 0.59 fewer days of out-of-school suspension ($p < 0.10$) and 2.91 fewer hours of in-school suspension than district comparison students ($p < 0.01$).

4. Exploratory Secondary Outcome Results

RQ4. *Did students who participated in the BellXcel Summer Program demonstrate higher self-confidence than matched comparison students?*

Using student responses to the OnTrack Greenville Student Survey, researchers assessed the impact of the BellXcel Summer Program on student self-confidence. Students in the treatment and district schools completed surveys in fall 2016 and fall 2017. Unlike the other outcomes, baseline data from before participating in the BellXcel Summer Program were not available for the survey; therefore, researchers were unable to look at changes after participation. More caution should be used when interpreting the results. However, the matching process included a number of variables that may have been related to the survey outcomes in order to mitigate the lack of baseline measures. For the analyses examining student self-confidence, positive values in the table indicate that the BellXcel scholars exhibited a more desirable outcome than the matched comparison students. The results for student self-confidence appear in Tables 28 and 29.

As shown in Table 28, students who completed the BellXcel Summer Program did not show significantly higher or lower levels of self-confidence (as measured by Academic Perseverance and Academic Self-Confidence) than matched comparison students at treatment and district schools in fall 2016.

Table 28. Exploratory Outcome Results for Student Self-Confidence Fall 2016

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size
Academic Perseverance	0.01 (0.09)	0.02	0.05 (0.10)	0.09
Academic Self-Confidence	0.04 (0.08)	0.08	0.04 (0.09)	0.06

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 324.
Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

In fall 2017, shown in Table 29, students who completed the BellXcel Summer Program had lower levels of self-confidence than matched comparison students at treatment and district, as measured by Academic Perseverance and Academic Self-Confidence. BellXcel scholars had significantly lower levels of Academic Self Confidence than matched students at treatment schools ($p < 0.05$) and district schools (p

< 0.05). Scholars also had lower levels of Academic Perseverance than matched students at district schools (p < 0.10).

Table 29. Exploratory Outcome Results for Student Self-Confidence Fall 2017

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched - Regression	Effect Size	Matched - Regression	Effect Size
Academic Perseverance	-0.12 (0.09)	-0.21	**-0.15†** (0.08)	-0.26
Academic Self-Confidence	**-0.19*** (0.09)	-0.34	**-0.21*** (0.08)	-0.38

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses. . BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.

Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

RQ5. *Did students who participated in the BellXcel Summer Program demonstrate a more positive attitude towards learning than matched comparison students?*

Researchers also examined the impact of the BellXcel Summer Program on student attitude toward learning through the OnTrack Greenville Student Survey. Again, baseline data from before participating in the BellXcel Summer Program was not available for this survey outcome measure. For the analyses examining student attitude toward learning, positive values in the table indicate that the BellXcel scholars exhibited a more desirable outcome than the matched comparison students. The results for Student Attitude toward Learning appear in Tables 30 and 31.

As shown in Table 30, BellXcel scholars did not have a significantly different attitude towards learning than matched comparison students at treatment or district schools in fall 2016. In addition, BellXcel scholars did not have a significantly different attitude towards learning than matched comparison students at treatment or district schools in fall 2017, shown in Table 31.

Table 30. Exploratory Outcome Results for Student Attitude toward Learning Fall 2016

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched - Regression	Effect Size	Matched - Regression	Effect Size
Valuing Education	-0.10 (0.08)	-0.21	-0.09 (0.09)	-0.19

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 324.

Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

Table 31. Exploratory Outcome Results for Student Attitude toward Learning Fall 2017

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched - Regression	Effect Size	Matched - Regression	Effect Size
Valuing Education	-0.07 (0.07)	-0.15	-0.01 (0.07)	-0.03

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

RQ6. *At the conclusion of the BellXcel Summer Program, did students who participated in the program demonstrate an increase in reading and math skills as measured by the STAR assessment scores?*

In addition to the confirmatory impact analyses that assessed student math and reading growth using the matching process, the research team also examined internal BellXcel program data to assess students' math and reading skills. BellXcel scholars completed the STAR assessment at the beginning and end of the program. While there is no comparison group for this analysis, one can examine how math and reading skills changed over the summer for BellXcel participants.

Examining data from summer 2016, BellXcel scholars scored lower on the STAR post-test than they did on the STAR pre-test. This is evident in Figure 4, which presents the reading pre-test score on the x-axis and the reading post-test scores on the y-axis. The majority of the observations are on or below the 45° line. A similar result is evident for the math exam, as shown in Appendix K.

To examine if the drop in scores is statistically significant, evaluators performed paired t-tests for both the reading and math exams in terms of scale scores (see Table 32). While BellXcel scholars scored 439 on the STAR reading exam at the beginning of summer, their scale scores decreased to 422 by the end of the BellXcel Summer Program. This decline was statistically significant. Over the course of the summer, the math scores decreased 8.1 points. However, this was not a statistically significant difference.

Figure 4. STAR Pre-test and Post-test Scale Scores in Reading Summer 2016

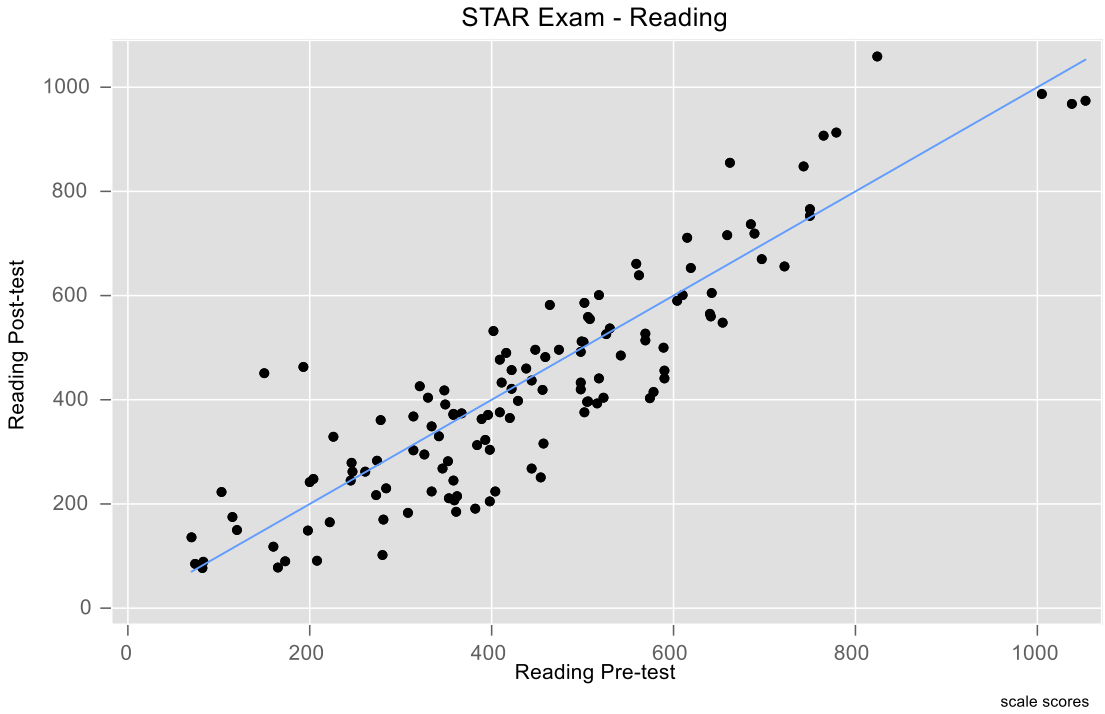


Table 32. STAR Pre-Test & Post-Test: Scale Scores Summer 2016

	Reading		**Math**	
	Mean	sd	Mean	sd
Post-Test	422.1	213.6	630.1	121.6
Pre-Test	438.5	193.6	638.2	115.0
Difference	**-16.4**†	95.6	-8.1	62.1
	N = 126, d = -0.08		N = 129, d = -0.07	

For the final STAR analysis of the summer 2016 scholars, researchers examined the relationship between BellXcel attendance and the change in STAR test scores from the beginning to the end of summer, while controlling for demographic factors. The researchers compared those scholars who attended at least 80% of the BellXcel program to those students who attended less than that threshold. Further, the relationship between the number of days attended and test score growth was also examined. The results in Table 33 indicate that the more days of BellXcel programming a scholar attended, the greater their reading and math test score growth (or there was less decrease). However, the relationship between attendance and math and reading growth were not statistically significant. Those that attended at least 80% of the BellXcel program scored higher on reading, but lower on math, than their counterparts, but the differences were not statistically significant.

Table 33. Days Attended BellXcel & Test Score Growth Summer 2016

	Reading STAR Scores	Reading STAR Scores	Math STAR Scores	Math STAR Scores
Attended 80% or More	20.7		-1.10	
	(21.9)		(13.2)	
# of Days Attended		0.42		0.26
		(2.79)		(1.86)
Pre-test	**0.96*****	**0.96*****	**0.87*****	**0.87*****
	(0.060)	(0.061)	(0.083)	(0.080)
Black	9.60	11.8	15.9	15.3
	(30.3)	(31.2)	(18.3)	(18.4)
Hispanic	-0.25	0.70	**47.6***	**47.1***
	(44.0)	(44.4)	(19.7)	(19.9)
Other Race	**118.3**$^{+}$	**113.0**$^{+}$	**60.4***	**62.2***
	(65.7)	(65.7)	(24.4)	(26.5)
Female	15.6	19.2	13.1	12.5
	(18.0)	(18.0)	(12.4)	(11.8)
Free/Reduced Lunch	-5.47	-4.67	3.33	3.92
	(36.0)	(36.4)	(13.7)	(14.5)
Special Ed.	**-39.3**$^{+}$	-37.4	-24.0	-24.5
	(23.4)	(23.5)	(17.3)	(17.0)
ESL	-35.0	-31.0	**-26.4**$^{+}$	**-27.0**$^{+}$
	(40.4)	(41.4)	(15.4)	(14.8)
Constant	-0.91	-1.12	55.9	50.9
	(56.5)	(80.2)	(57.8)	(64.9)
N	125	125	128	128
Adj. R^2	0.81	0.81	0.74	0.74
Cohen's D	0.10	0.03	-0.01	0.03

Robust SE in parentheses; † p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001. OLS Regression
Note: The Cohen's d estimate reflects the effect size of the "Attend 80% or More" or the "# of Days Attended" variable for each regression.

A similar analysis was performed using pre-test and post-test data from scholars who participated in the BellXcel program in summer 2017. A simple comparison of reading scores indicates that BellXcel scholars improved their performance over the summer, as seen in Figure 5. See the scatterplot for math achievement in Appendix K. Table 34 presents paired t-tests for the reading and math test scores. Examining those students with complete data, one sees that BellXcel scholars increased their performance on the reading post-test by approximately 23 points, on average. This was a statistically significant difference ($p < 0.01$). BellXcel scholars also exhibited significant improvement in their math outcomes ($p < 0.10$).

Figure 5. STAR Pre-test and Post-test Scale Scores in Reading Summer 2017

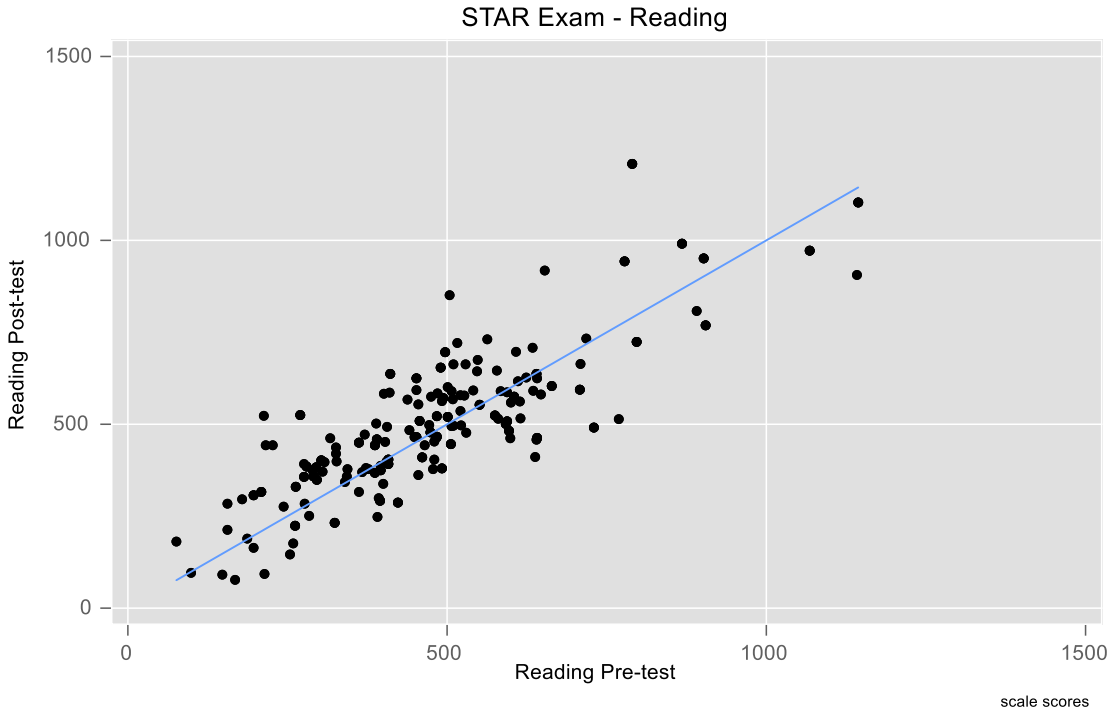


Table 34. STAR Pre-Test & Post-Test: Scale Scores Summer 2017

	Reading		**Math**	
	Mean	sd	Mean	sd
Post-Test	502.3	198.5	673.9	96.4
Pre-Test	479.3	196.8	666.4	97.6
Difference	**22.9****	112.9	**7.5**$^{+}$	61.2
	N = 209, d = 0.12		N = 209, d = 0.08	

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Researchers also estimated OLS regressions predicting math and reading post-test scale scores while controlling for student demographics. Table 35 indicates that BellXcel scholars who attended at least 80% of the summer program scored significantly higher on the math post-test than those who attended less than 80% ($p < 0.05$). The Cohen's d value was 0.23. While the other regressions indicated that those with higher BellXcel attendance demonstrated higher reading and math achievement, the relationships were not statistically significant.

Table 35. Days Attended BellXcel & Test Score Growth Summer 2017

	Reading STAR Scores	Reading STAR Scores	Math STAR Scores	Math STAR Scores
Attended 80% or More	11.1 (17.4)		**20.5*** (9.13)	
# of Days Attended		1.80 (1.37)		1.10 (0.76)
Pre-test	**0.85***** (0.049)	**0.85***** (0.047)	**0.74***** (0.056)	**0.75***** (0.056)
Black	7.54 (27.5)	4.38 (26.9)	-6.18 (13.2)	-7.75 (13.6)
Hispanic	-28.6 (25.5)	-31.1 (24.9)	2.32 (14.7)	4.07 (14.8)
Other Race	-49.7 (36.6)	-56.8 (35.9)	-22.0 (22.9)	-28.5 (22.5)
Female	6.43 (18.2)	6.93 (18.1)	-2.73 (8.38)	-2.25 (8.45)
Free/Reduced Lunch	-3.37 (19.6)	-6.55 (19.5)	-15.3 (10.7)	-13.8 (11.2)
Special Ed.	-11.1 (20.5)	-12.4 (20.5)	**-24.7**$^{+}$ (12.9)	**-23.3**$^{+}$ (12.8)
ESL	23.9 (22.1)	22.4 (22.1)	-6.58 (12.5)	-8.86 (12.4)
Constant	96.6* (40.8)	77.2^{+} (44.7)	189.6*** (38.2)	179.3*** (38.9)
N	175	175	176	176
Adj. R^2	0.70	0.71	0.65	0.65
Cohen's d	0.06	0.10	0.23	0.25

Robust SE in parentheses; † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. OLS Regression
Note: The Cohen's d estimate reflects the effect size of the "Attend 80% or More" or the "# of Days Attended" variable for each regression.

RQ7. *At the conclusion of the BellXcel Summer Program, did the parents of students who participated demonstrate an increase in engagement in their children's education?*

Researchers measured this outcome through a single-item, self-report survey question located on the post-only BellXcel Parent Survey. There was no comparison group for this analysis. In summer 2016, parent respondents (N = 43) overwhelming indicated that they were more involved in their child's learning since enrolling their child in the BellXcel Summer Program. Among respondents, 93% of parents agreed that they were more involved in their child's learning. Only 2% of parents disagreed that they were more involved, and 5% were not sure. In summer 2017, parent respondents (N = 98) still indicated that they were more involved in their child's learning, though there was a decrease from the prior summer. Only 78% of parents agreed that they were more involved, 13% disagreed that they were more involved, and 9% of parents were not sure.

5. Additional Exploratory Outcomes

RQ8. *Did students who participated in the BellXcel Summer Program report improved relationships with their teachers?*

RQ9. *Did students who participated in the BellXcel Summer Program report improved relationships with caring adults in their school?*

Researchers examined the impact of the BellXcel Summer Program on student relationships with caring adults. Baseline data from before participating in the BellXcel Summer Program were not available for this outcome measure from the OnTrack Greenville Student Survey. For the analyses examining student relationships with caring adults, positive values in the table indicate that the BellXcel scholars exhibited a more desirable outcome than the matched comparison students. The results for student relationships with caring adults appear in Tables 36 and 37. Scholars who participated in the BellXcel Summer Program in summer 2016 or summer 2017 did not report significantly stronger relationships with teachers and caring adults at their schools when compared to matched students treatment or district schools.

Table 36. Exploratory Outcome Results for Relationships with Caring Adults Fall 2016

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size
Relationships with Teachers	-0.12 (0.10)	-0.20	0.04 (0.12)	0.06
Relationships with Caring Adults	-0.08 (0.09)	-0.13	0.09 (0.11)	0.13

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 324.
Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

Table 37. Exploratory Outcome Results for Relationships with Caring Adults Fall 2017

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size
Relationships with Teachers	-0.14 (0.10)	-0.21	0.03 (0.08)	0.06
Relationships with Caring Adults	-0.13 (0.10)	-0.22	0.05 (0.09)	0.07

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

RQ10. *Did students who participated in the BellXcel Summer Program report improved school engagement?*

Researchers also assessed the impact of the BellXcel Summer Program on student engagement at school. Two measures of school engagement, School Engagement and School Belonging, were included on a post-program survey administered in fall 2016 and 2017. Baseline data from before participating in the BellXcel Summer Program were not available for this outcome measure from the OnTrack Greenville Student Survey. For the analyses examining student engagement at school, positive values in the table indicate that BellXcel scholars exhibited a more desirable outcome than matched comparison students. The results for student engagement at school appear in Tables 38 and 39. Overall, BellXcel scholars did not report significantly higher or lower levels of School Engagement or School Belonging than matched comparison students at treatment or district schools in fall 2016 or fall 2017. The only exception is that BellXcel scholars reported significantly lower levels of School Belonging in fall 2017 when compared to matched students at treatment schools ($p < 0.10$).

Table 38. Exploratory Outcome Results for School Engagement Fall 2016

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Standard Mean Difference	Matched Regression	Standard Mean Difference
School Engagement	-0.15 (0.10)	-0.23	-0.08 (0.12)	-0.12
School Belonging	-0.02 (0.09)	-0.03	0.05 (0.09)	0.07

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 324.
Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

Table 39. Exploratory Outcome Results for School Engagement Fall 2017

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size
School Engagement	-0.10 (0.10)	-0.15	-0.03 (0.09)	-0.05
School Belonging	**-0.18†** (0.10)	-0.27	-0.07 (0.09)	-0.09

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

RQ11. *Did students who participated in the BellXcel Summer Program report participating in educational activities over the summer at a higher frequency than comparison students?*

Lastly, researchers examined self-reported student behaviors over the summers of 2016 and 2017 to assess if participating in BellXcel changed the types of activities that students were involved in over the summer. For the analyses examining the frequency of summer activities, positive values in the table indicate that the BellXcel scholars exhibited more frequently completed the activity. Results for the frequency of summer activities appear in Tables 40 and 41.

As shown in Table 40, there were few significant differences in the frequency with which BellXcel scholars participated in certain summer activities when compared with matched students at treatment schools and district schools for the summer of 2016. BellXcel scholars did report playing more math games ($p < 0.05$) than matched students at district schools over the summer of 2016.

Table 40. Exploratory Outcome Results for Frequency of Summer Activities Fall 2016

How often did you…	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size
…go to the library?	0.16 (0.34)	0.09	0.14 (0.37)	0.08
…write something?	0.18 (0.29)	0.10	0.05 (0.30)	0.03
…play math games?	0.27 (0.30)	0.15	**0.97*** (0.33)	0.53
…read a book?	-0.16 (0.27)	-0.09	-0.24 (0.29)	-0.14
…play on your phone, watch TV, or play video games?	0.27 (0.34)	0.15	-0.35 (0.42)	-0.20
…go to community center, YMCA, or camp?	-0.04 (0.30)	-0.02	-0.29 (0.30)	-0.16
…play outside?	0.33 (0.31)	0.18	0.03 (0.29)	0.02

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 324.
Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

There were several significant differences in the frequency with which BellXcel scholars participated in educational activities over the summer of 2017 when compared with matched students at treatment schools and district schools. See Table 41. When compared with matched students at treatment schools, BellXcel scholars more frequently reported writing ($p < 0.001$) and going to a community center, YMCA, or camp ($p < 0.10$) during the summer of 2017. When compared to matched students at district schools, BellXcel scholars more frequently reported writing ($p < 0.01$), playing math games ($p < 0.05$), and going to a community center, YMCA, or camp ($p < 0.05$) during the summer of 2017. Additionally, BellXcel

scholars less frequently reported spending time playing on their phone, watching TV, or playing video games than matched students at district schools ($p < 0.05$).

Table 41. Exploratory Outcome Results for Frequency of Summer Activities Fall 2017

How often did you…	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Matched Regression	Effect Size	Matched Regression	Effect Size
…go to the library?	0.23 (0.29)	0.13	0.14 (0.27)	0.08
…write something?	**1.08*** ** (0.30)	0.59	**0.92** ** (0.30)	0.51
…play math games?	0.35 (0.29)	0.19	**0.61*** (0.30)	0.34
…read a book?	0.46 (0.30)	0.26	0.32 (0.27)	0.17
…play on your phone, watch TV, or play video games?	-0.28 (0.34)	-0.16	**-0.66*** (0.28)	-0.36
…go to community center, YMCA, or camp?	**0.58†** (0.31)	0.32	**0.54*** (0.25)	0.30
…play outside?	0.07 (0.33)	0.04	-0.25 (0.27)	-0.14

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and bootstrap SEs in parentheses from multivariate regressions that also control for student demographics and grade. Frequency weights were used to account for matching with replacement.

6. Adjustment for Multiple Outcomes

This evaluation examined the effects of completing the BellXcel Summer Program on multiple outcomes. However, as the number of comparisons in a study increases, so does the possibility of committing a Type I error. There are many ways to deal with this challenge. As seen above in section II, this evaluation examined outcomes in a number of different domains. Following Schochet (2008), the impacts and outcomes of interest were classified as confirmatory or exploratory. The confirmatory analyses in course performance, as measured by standardized test scores, were the focus of this evaluation. Before adjustments for multiple comparisons were made, none of the confirmatory tests found statistically significant differences between BellXcel and comparison students; therefore, it was not necessary to adjust for multiple outcomes for confirmatory analyses in course performance.

Comparisons between BellXcel and comparison students on behavior and attendance were classified as exploratory outcomes. Within each of these domains, the Benjamni-Hochberg adjustment was used. Researchers ordered the p-values from the tests from smallest to largest and compared each to an adjusted p-value that took the number of tests in the domain into account.

The multiple comparison adjustment for the behavior domain is presented below in Table 42. There were 30 total tests (any referral, any in-school suspension, any out-of-school suspension, number of referrals, number of hours of in-school suspension, and number of days of out-of-school suspension for the five matches using the administrative data). Before adjustments for multiple comparisons were made, nine of the tests indicated statistically significant differences between BellXcel scholars and comparison students. All of the results indicated better levels of discipline among BellXcel scholars. After adjusting for multiple comparisons, one result remained statistically significant.

Table 42. Multiple Comparison Adjustment for the Behavior Domain

Outcome	Year	Comparison Group	Coefficient	Original p-value	Significant at p < 0.10
# Hours ISS	2017	District Schools	-5.24	0.00146	Yes
# Days OSS	2017	District Schools	-0.96	0.00881	No
# Days OSS	2018	Treatment Schools	-0.73	0.00955	No
# Hours ISS	2018	District Schools	-2.91	0.01002	No
# Behavioral Referrals	2017	District Schools	-0.64	0.01873	No
Any OSS	2018	Treatment Schools	-0.82	0.02656	No
Any ISS	2017	District Schools	-0.92	0.03130	No
# Hours ISS	2017	State	-0.26	0.05967	No
# Days OSS	2018	District Schools	-0.59	0.07165	No

The multiple comparison adjustment for the attendance domain is presented below in Table 43. There were ten total tests (average daily attendance rate and chronic absenteeism for the five matches using the administrative data). Before adjustments for multiple comparisons were made, two of the tests indicated that BellXcel scholars demonstrated statistically significantly higher average daily attendance than comparison students. After adjusting for multiple comparisons, both results remained statistically significant.

Table 43. Multiple Comparison Adjustment for the Attendance Domain

Outcome	Year	Comparison Group	Coefficient	Original P-value	Significant at p < 0.10
Average Daily Attendance	2018	Treatment Schools	0.977	0.00510	Yes
Average Daily Attendance	2018	District Schools	0.943	0.00602	Yes

V. Findings, Lessons Learned, and Next Steps

A. Summary of Implementation Study Results

Based on collection and analysis of data from observations, interviews, focus groups, and surveys of staff, scholars, and parents, researchers concluded that the BellXcel Summer Program was implemented in treatment middle schools with a relatively high degree of fidelity to the BellXcel logic model. Further, the commitment of the staff, including the site program managers, program assistants, teachers, and assistant teachers, observed by researchers during site visits was notable. Summer academic programs can be challenging to implement, as students can be resistant to "going to school" in the summer and teachers of those programs have less time to decompress from the previous school year and rest for the next school year. Nevertheless, the atmosphere of the BellXcel Summer Program at all three sites was positive and productive. It was clear that teachers took their job of providing scholars with summer instructional time seriously. Teachers also generally felt well supported by BellXcel Carolinas staff and had relatively high levels of satisfaction with the program.

One hundred percent of BellXcel staff at the three sites completed the in-person BellXcel training in summer 2016, while 87% of staff at School 3, 64% at School 4, and 50% at School 1 completed the additional online training in summer 2017. Teachers (including lead teachers, assistant teachers, and instructional coaches) were overall satisfied with the extent to which the training prepared them for their roles.

A key input of the BellXcel logic model is a curriculum aligned with the Common Core. BellXcel teachers reported implementing the BellXcel curriculum in both ELA and math, though many teachers indicated they modified or supplemented the curriculum to meet the needs of their scholars, many of whom were performing substantially below grade level. Veteran teachers noted improvements in the BellXcel math curriculum by broadening the focus beyond fractions.

In addition to ELA and math, all three sites implemented an enrichment curriculum. During a site visit, the researchers observed enrichment classes in dance, social skills, technology, cooking, art, and karate. In every case, scholars had access to appropriate and interesting materials and appeared engaged and joyful. In most of these classes, teachers drew connections to careers that aligned to the activity (e.g., chef and chemist for cooking, engineer and graphic designer for art).

In summer 2017, community time, another curricular element designed to start the day with a fun, interactive, and structured activity, was executed well across the three sites. This is an improvement from summer 2016, when only one site appeared to implement community time with fidelity.

Teachers generally reported positive views on behavior management, with 100% of teachers agreeing or strongly agreeing that the behavior management system allowed for fair and respectful treatment of scholars and that it allowed for scholars to learn self-management. Scholars were slightly less positive about scholar behavior. For example, in summer 2017, 38% of School 4 scholars reported that scholars almost always behaved well, compared with 35% at School and 24% at School 3.

The BellXcel logic model specifies a goal that program sites will have an 80% or higher average daily attendance. Both School 3 and School 4 achieved average daily attendance of at least 80% in summer 2017 (82% and 81%, respectively), and School 3 came very close at 78%.

In summer 2017, very high percentages of scholars reported enjoying the field trips (92%) and enrichment classes (85%). Meanwhile, over half of the scholars (59%) reported that they liked the academic classes. Notably, while scholars at the three sites were similarly likely to report finding their math classes extremely or quite interesting, the extent to which they found ELA interesting varied. At School 3, 81% of scholars found ELA extremely or quite interesting, compared with 78% at School 1 and 52% at School 4. Scholars across all three sites were equally likely to indicate they enjoyed coming to BellXcel. However, scholars had slightly less favorable feedback about whether they would tell their friends to come to BellXcel (overall, 60% said they would) and whether they would go to BellXcel again (overall, 65% said they would).

Parent engagement events are an important part of the BellXcel Summer Program. The logic model specifies a goal of 70% of parents attending parent engagement events. Actual participation rates varied from as low as 7% to as high as 51%.

B. Summary of Impact Study Results

This study targeted a moderate level of evidence and utilized a single-site non-randomized group design with groups formed by propensity score matching. For confirmatory impact research questions, there were three comparison groups. Treatment students were matched to (1) other students in the *treatment schools* who did not participate in the intervention; (2) other students in the *same school district* attending district schools; and (3) other students attending *Title I schools across the state* of South Carolina. The use of these multiple comparison groups improved the overall internal and external validity of the study, as each comparison group presented different threats to validity. Researchers matched students using a propensity score model that included race, gender, grade level, English proficiency, special education status, free and reduced meal eligibility, and baseline outcome variables. Researchers conducted separate matching procedures for each data source, administrative data and survey data. At the conclusion of the matching process, researchers ensured that there were no significant differences between the treatment and comparison groups on pre-treatment covariates.

Confirmatory Impact Results

This study examined the impact of participation in the BellXcel Summer Program on student course performance in math and ELA, as measured by growth in MAP math and reading assessment scores. Researchers were not able to confirm the hypothesis that BellXcel scholars would have improved course performance when compared to matched comparison students. In summer 2016, there were no statistical differences in math or ELA course performance as measured by growth in MAP reading and math test scores when comparing BellXcel scholars to matched comparison students at treatment and district schools. Due to an unexpected change in the district testing schedule, MAP assessment scores were not available for academic year 2017-18; therefore, researchers were not able to test the impact of participation in BellXcel in summer 2017.

Exploratory Impact Results - Attendance

Researchers examined the impact of participating in the BellXcel Summer Program on student attendance, confirming the hypothesis that BellXcel students would have improved attendance when compared to matched comparison students. Summer 2017 BellXcel scholars had higher average daily attendance than their matched counterparts attending treatment ($p < 0.01$, $d = 0.27$) and district schools ($p < 0.01$, $d = 0.25$). There were no statistically significant differences in chronic absenteeism between BellXcel scholars and matched comparison students during either academic year.

For the exploratory impact analyses examining student attendance during the academic year following program participation, there were ten total tests (average daily attendance rate and chronic absenteeism for the five matches using the administrative data). Before adjustments for multiple comparisons were made, two of the tests indicated that BellXcel scholars demonstrated statistically significantly higher average daily attendance than comparison students. After adjusting for multiple comparisons, both results remained statistically significant.

Exploratory Impact Results - Behavior

Researchers examined the impact of participation in the BellXcel Summer Program on student behavior. Results varied greatly by academic year, comparison group, and type of behavioral incident, but results generally showed that BellXcel scholars had fewer behavioral incidences than matched comparison students.

Summer 2016 BellXcel scholars had fewer behavioral referrals during the academic year following program participation than matched comparison students attending district schools ($p < 0.10$. $d = -0.19$). No other differences in behavioral referrals between treatment and comparison students were statistically significant.

Across both years of the study, BellXcel scholars generally had significantly fewer days of out-of-school suspension than matched comparison students. Summer 2016 BellXcel scholars received 0.96 fewer days of out-of-school suspension that matched students at district schools ($p < 0.01$, $d = -0.26$). Summer 2017 BellXcel scholars received 0.73 fewer days of out-of-school suspension than matched students at treatment schools ($p < 0.05$, $d = -0.23$) and 0.59 fewer days than matched students at district schools ($p < 0.05$, $d = -0.18$). Further, summer 2017 BellXcel scholars were less likely to have received any out-of-school suspension than matched students at treatment schools ($p < 0.05$, $d = -0.45$).

BellXcel scholars generally had fewer hours of in-school suspension than matched students at district schools. Summer 2016 BellXcel scholars had 0.72 fewer hours of in-school suspension ($p < 0.10$, $d = -0.41$) and summer 2017 BellXcel scholars had 2.91 ($p < 0.01$, $d = -0.23$) fewer hours of in-school suspension than their matched counterparts at district schools. In addition. Summer 2016 BellXcel scholars were less likely to have received any in-school suspension than matched comparison students at district schools ($p < 0.01$, $d = -0.51$) and received fewer total in-school suspensions than matched comparison students attending state schools ($p < 0.10$, $d = -0.16$).

For the exploratory impact analyses examining student behavior during the academic year following program participation, there were 30 total tests (any referral, any in-school suspension, any out-of-school suspension, number of referrals, number of hours of in-school suspension, and number of days of out-of-school suspension for the five matches using the administrative data). Before adjustments for multiple comparisons were made, nine of the tests indicated statistically significant differences

between BellXcel scholars and comparison students in the predicted direction. After adjusting for multiple comparisons, one result remained statistically significant.

Exploratory Secondary Outcome Results

The secondary goals of the BellXcel Summer Program are to help students improve their educational self-confidence, attitude toward learning, and growth in reading and math skills. The program also aims to improve parents' involvement in their children's learning. An analysis of student survey data, an internal program learning assessment, and post-program parent survey data allowed researchers to assess the effect of participation in the BellXcel Summer Program on these exploratory secondary outcomes.

When compared to matched comparison students, summer 2016 BellXcel scholars did not have higher levels of self-confidence or an improved attitude toward learning after participating in the BellXcel Summer Program. In summer 2017, BellXcel scholars had lower levels of self-confidence as matched comparison students at district schools, as measured by Academic Self-Confidence ($p < 0.05$, -0.26) and Academic Perseverance ($p < 0.10$, -0.38).

Scholar growth in math and reading varied by program year, as measured by an internal program assessment. In summer 2016, BellXcel scholars did not show any growth in math or reading skills over the course of the program in summer 2016. In Summer 2017, however, BellXcel scholars did demonstrate significant growth in math ($p < 0.10$, $d = 0.08$) and reading skills ($p < 0.01$, $d = 0.12$).

The majority of parents of BellXcel scholars reported on a post-program survey that they were more involved in their student's learning since enrolling their student in the BellXcel Summer Program. In summer 2016, 93% of parents reported that they were more involved in their child's learning. In summer 2017, 78% of parents reported that they were more involved.

Additional Exploratory Results

The OnTrack Greenville Student Survey included additional survey measures used to test secondary exploratory outcomes of other Sub-Grantee interventions in the OnTrack Greenville SIF portfolio. While these outcomes did not appear in the BellXcel Summer Program logic model, researchers assess the impact of participation in the program on student self-reported summer activities, student engagement at school, and student relationships with adults at school.

When reflecting on their summer activities on a post-program survey, Summer 2016 BellXcel scholars more frequently reported playing math games over the summer than matched comparison students attending district schools ($p < 0.01$, $d = 0.53$). There were no other significant differences in self-reported summer activities between scholars and matched comparison students in treatment and district schools in summer 2016. Researchers hypothesized that treatment students would report more frequently going to a community center or camp during the summer month than matched students. The lack of significant results introduces the possibility that matched comparison students participated in summer learning experiences that researchers were not able to account for in the matching process.

The following year, summer 2017 BellXcel scholars more frequently reported writing ($p < 0.01$, $d = 0.51$), playing math games ($p < 0.05$, $d = 0.34$), and going to a community center or camp ($p < 0.10$, $d = 0.30$) over the summer than matched students at district schools. They also reported less frequently playing on a phone,

watching TV, or playing video games than district school matches ($p < 0.05$, $d = -0.36$). In addition, BellXcel scholars were more frequently reported writing ($p < 0.05$, $d = 0.59$) and going to a community center or camp ($p < 0.10$, $d = 0.32$) than matched students at treatment schools.

In general, there were no significant differences in school engagement or relationships with adults at school between BellXcel scholars and matched comparison students at either school group. The only exception is that summer 2017 BellXcel scholars did report lower levels of school belonging than matched students attending treatment schools ($p < 0.10$, $d = -0.27$).

Evidence Level Determination and Discussion of Results

Results of the analyses do not yet provide evidence supporting the confirmatory impact research question. After participating in the BellXcel Summer Program, treatment students did not show improved course performance in math or reading when compared to matched students, as measured by growth in math and reading MAP assessment scores. Data from an internal pre/post-program assessment of growth in math and reading skills showed that BellXcel scholars experienced significant growth in math and reading skills in summer 2017. Data from this program assessment, the STAR assessment, were not available for matched comparison students. If MAP assessment scores had been available in academic year 2017-18, it is possible that the significant growth in math and reading skills shown on the STAR assessment also would have translated to significant growth in math and reading MAP scores when compared to matched comparison students. In the final two years of the study, researchers will explore other opportunities for measuring the math and ELA course performance of treatment and comparison students.

Exploratory impact analyses did reveal significant effects of the program in attendance and behavior, even after adjusting for multiple outcomes, allowing the study to achieve a moderate level of evidence. Though not intended outcomes of the program, exploratory analyses found that BellXcel scholars had higher average daily attendance and fewer hours of in-school suspension than matched comparison students, with results varying by year and comparison group. Additional analyses are needed to understand better the relationship between participation in the BellXcel Summer Program and positive effects in attendance and behavior.

One possible explanation for these positive results in attendance and behavior relates to the types of variables available for use in the propensity score matching process and that the administrative data used in the matching process did not fully capture the propensity for participating in the BellXcel Summer Program. For example, this study included in the treatment group scholars who attended 80% or more of the voluntary summer program. While the propensity score matching process included pre-treatment outcome measures, such as prior year attendance and behavior, the transition from fifth grade to sixth grade poses many challenges and sixth-grade students' attendance and behavior typically worse when compared to their elementary school outcomes. It is possible students who were more likely to attend the BellXcel Summer Program with a high rate of attendance also were more likely to have a high rate of attendance and good behavior in sixth grade and that the combination of demographic variables and prior year outcomes did not measure this likelihood.

The secondary goal of the BellXcel Summer Program is to help students improve their academic self-confidence and attitude toward learning. In this first year of the impact study, there were no significant differences between BellXcel scholars and matched comparison students at treatment or district schools in the secondary outcomes, which were measured on a post-program student survey. Scholars who

participated in summer 2017 had lower levels of self-confidence than matched comparison students at treatment and district schools. It is possible, however, that scholars who participated in the BellXcel Summer Program were more aware of concepts related to self-confidence through their participation in the program, which led to a higher degree of self-reflection and lower responses to these survey items.

When examining other additional exploratory outcomes from the student survey, researchers did identify positive exploratory findings. When asked about the frequency with which they participated in certain activities over the summer, 2017 BellXcel scholars more frequently reported playing math games, writing, and going to a community center or camp over the summer than matched comparison students attending district schools. BellXcel scholars also reported less frequent use of a device (phone, TV, or video games) over the summer months than matched comparison students at district schools. These findings indicate that BellXcel scholars were using their time over the summer in a substantially different way than matched students, which presumably could lead to academic or social-emotional gains in future years of the study.

The absence of significant positive findings in course performance for in-school matches may be related to missing data on student participation in other OnTrack Greenville interventions. OnTrack Greenville includes four other formal implementation partners and several informal partners, some of whom are working to improve the same student outcomes as the BellXcel Summer Program. It is possible that some of the in-school matches selected for the present study participated in other OnTrack Greenville support programs that influenced student course performance. Future analyses will include more complete data on participation in other OnTrack Greenville interventions, allowing researchers to control for participation during the in-school matching process. While researchers will not have access to comparable participation data for students in district schools, none of the OnTrack Greenville programs operate in comparison schools and there are few support programs similar to OnTrack Greenville available to students at these schools. Therefore, there is less concern that comparison students at district schools also are participating in potentially similar programs.

Moving forward, researchers will consider other measures for course performance in future years of the study. As discussed throughout this report, in academic year 2017-18, the local school district discontinued the use of the MAP assessment, replacing it with a different computerized adaptive test that is administered three times per year, Mastery Connect. There is more to learn about Mastery Connect as a possible measure for course performance and its utility for the study. Further, this change in outcome measure would prohibit researchers from comparing some test results from year to year. Other studies (e.g. Corrin et al., 2016) have opted to include measures of course grades, examining the number of core courses failed, instead of standardized test scores. While researchers only would have access to course grade data for students within the local school district and not state-wide, this is an additional option for refining the impact analysis of course performance moving forward.

Changes to the SEP - Impact Study Design

The primary change to the SEP was the loss of a key outcome measure of course performance, MAP assessment scores in ELA and math. Prior to academic year 2017-18, the local school district administered the MAP assessment in grades 3 through 8 at least two times per year, in the fall and spring. Some schools opted to administer the assessment a third time, in winter. The district opted to end its contract with MAP and began administering Mastery Connect in fall of 2017. At present time, researchers do not have access to Mastery Connect data and remain uncertain if data from this assessment will serve as an acceptable outcome measure in the study. Researchers only were able to

access end-of-year SC READY assessment scores in math and ELA for academic year 2017-18. Using this end-of-year assessment as an outcome measure for the evaluation of a summer program is problematic, however, as roughly ten months have passed since students received the intervention and the administration of the standardized test. This period of time separating treatment and assessment is far too long, as students experience an entire academic year of learning and any differences in test scores could not reasonably be attributed to participation in the BellXcel Summer Program. As such, researchers did not include SC READY as an outcome measure for confirmatory impact questions. This meant that researchers only were able to test confirmatory research hypotheses for students who attended the program in summer 2016.

Researchers also had to alter the inclusion criteria for state comparison schools. In the SEP, researchers originally proposed to select state schools with a poverty index of 85 or higher and a Hispanic student population of at least 10%. At the time of writing the SEP, South Carolina calculated the poverty index based on the number of students eligible for free or reduced meals. After the introduction of the community provision for free and reduced meals, state officials introduced a new measure of poverty that included students who met any of the criteria: homeless or migrant during the academic year; Medicaid enrollment at any time during a three year period; SNAP enrollment at any time during a three year period; TANF enrollment at any time during a three year period; or foster care enrollment at any time during a three year period. This change in the poverty index affected and in general reduced the reported poverty levels of treatment schools and all schools across the state. When researchers searched for state comparison schools using these two criteria, only 13 schools appeared as possible comparison schools from which to draw matched comparison students. Of these schools, several were charter schools or schools serving exceptional learners and were not appropriate to serve as comparison schools. As such, researchers relaxed the inclusion criteria for state comparison schools and included all Title I middle schools in the state of South Carolina outside of the local district.

In addition, researchers were not certain if they would be able to administer the OnTrack Greenville Student Survey outside of the treatment schools. The local district allowed researchers to administer this pre- and post-survey at the four district comparison schools, allowing researchers to analyze student survey outcomes using matched comparison students. This change to the SEP strengthened the study's design for the analysis of secondary research questions.

C. Lessons Learned

Program staff have learned valuable lessons around program recruitment, enrollment, and attendance. As program implementation has progressed in OnTrack Greenville, each year has presented different challenges and learning opportunities. In the early years, there were issues with recruiting and enrolling the target population. Program staff learned that, even as last as February of 2017, there were feeder elementary school principals who were not fully aware of the OnTrack Greenville initiative and/or of BellXcel's role as an OnTrack Greenville partner. In early 2017, enrollment efforts improved, with the program operating with a waiting list going into summer 2017. Program staff attributed this enrollment increase to a successful grassroots effort, which involved attending local parent engagement events, as well as canvassing surrounding neighborhoods. However, many enrolled students did not actually attend the program, causing program staff to move through the wait list quickly and still having less than expected program attendance. Reasons students did attend included moving out of the district or state before the summer program started and competing summer opportunities, such as spending part of the

summer with a relative. Staff have started to try different approaches to emphasizing the attendance commitment required when enrolling students to the program. Program staff also noticed a significant drop in attendance the week of the Fourth of July, with many students missing part or all of the week. As such, leaders decided to close the program that week in summer 2018. Finally, BellXcel staff also learned that parents are not ready or able to fully embrace online tools, and that paper applications and newsletters are still the best format for communication with parents and families.

Lessons learned regarding the implementation of BellXcel's program include the need for an increased effort surrounding community sponsorships. Program staff acknowledge that more in-kind support from community partners could likely be generated, ultimately leading to increased resources available to students. In addition, ensuring that BellXcel's local leadership is in place to fully engage with OnTrack and community partners has ensured a smoother implementation of both the recruitment and program implementation processes. Full transparency about methods and structure of BellXcel's program also has emerged as an important component of success. In regards to curriculum components, BellXcel now incorporates PEAR, which is a research-based assessment of Social-Emotional learning. Program staff believe that, in addition to providing academic instruction, addressing social-emotional skills is critical to creating well-rounded students who are poised for success.

BellXcel stakeholders also continue to learn about scaling the program to serve students in upper grade levels. In summer 2016 and 2017, BellXcel only served rising sixth-grade students. As BellXcel analyzed attendance data, they learned more about which students excelled in the program and maintained high attendance, program staff worked with schools to discuss scaling plans. In summer 2018, not included in the present impact study, BellXcel began to serve a small number of rising seventh-grade students. These students, identified in partnership with OnTrack Greenville schools, already attended the BellXcel Summer Program as sixth graders with satisfactory attendance, showed need for ongoing assistance in math and ELA, and were excited about returning to the program. Future evaluation efforts will explore the impact of this repeated exposure to the BellXcel Summer Program.

There were several lessons learned related to the evaluation itself. First, there were lessons learned about the importance of data-sharing agreements and school-community partnerships. OnTrack Greenville is a collective impact initiative that created a culture of trust and learning with Sub-Grantee and school partners. Among the six formal guiding values of the partnership, two values helped foster a strong culture that supported the evaluation: (1) operating as an innovative learning community and (2) having a results-oriented mindset. Partners' commitment to these shared values and the ongoing efforts of United Way of Greenville County serving as the collective impact backbone helped strengthen relationships between researchers, district stakeholders, and partners. For that reason, partners over time grew more comfortable with evaluation and embraced learning opportunities rather than fearing potentially negative or unexpected findings. This helped to ensure that implementation and impact study findings and recommendations more quickly translated into programmatic changes. Without the trust and shared values of the OnTrack Greenville partnership, there would have been additional challenges threatening the validity of the study.

D. Study Limitations

There were many strengths to this study. The use of multiple comparison groups improved the overall internal and external validity of the study, as each comparison group presented different threats to validity. The majority of positive significant findings were detected with the matched students attending district comparison schools. One strength is that these schools shared the same district and community context. Moreover, students in this comparison group were likely to have participated in the BellXcel Summer Program if it had been available to them at their school. These schools did not share the same school or neighborhood contexts, though, presenting a threat to internal validity.

In addition, a thorough implementation study strengthened the implementation of the program and allowed researchers to confirm a sufficient degree of model fidelity. The lessons learned through the implementation study were valuable to project stakeholders and helped shine a light on program strengths and possible areas of improvement.

However, there were several limitations to the study. First, researchers were not able to identify a subset of state Title I middle schools with student population demographics similar to the treatment schools. The Sub-Grantee Evaluation Plan called for including only state comparison schools with a poverty index of 85% or higher and Hispanic students representing 10% of the student body. Only 13 schools met these inclusion criteria to be considered as state comparison schools—many were charter schools or special designation schools serving students with disabilities and were substantially different from the treatment schools. Researchers opted to loosen the inclusion criteria and include the approximately 300 Title I middle schools in South Carolina in the state school comparison group.

Another limitation of the study is that researchers did not have the ability to assess if comparison students at district and state schools had received similar program services, such as other summer learning programs. While researchers could confirm that state comparison schools did not offer the exact BellXcel Summer Program, it is possible that a number of schools in the state comparison group offered similar summer programs for rising sixth-grade students. However, the number of matched students who attended these schools likely was very small and the inclusion of these students as matches would not have influenced the results of the study significantly.

A major limitation relates to the lack of pre- and post-program measures of course performance widely available for treatment and comparison students. When the study began, researchers planned to use the MAP assessment to measure course performance, as students took the assessment in the spring before the summer program and again in the fall following the summer program. Though this was not a perfect outcome measure, it was the best fit for the study at the time. In academic year 2017-18, Greenville County Schools discontinued the administration of the MAP assessment for most grade levels. Without access to MAP assessment data, researchers did not have access to a suitable measure for the confirmatory impact research question.

In addition, the treatment schools simultaneously were implementing formal and informal school-wide initiatives to improve student attendance, behavior, and course performance. These school-wide efforts were confounding factors that may explain the lack of significant effects when comparing BellXcel students to in-school matched comparison students. These school-wide efforts also increased the likelihood that the positive significant effects of the program identified when examining district school matches may not be fully attributable to the BellXcel Summer Program

Further, the absence of positive significant findings for in-school matches may be related to missing data on student participation in other OnTrack Greenville interventions. Apart from the school-wide models discussed above, OnTrack Greenville includes four other formal implementation partners and several informal partners, some of whom are working to improve the same student outcomes as the BellXcel Summer Program. It is possible that some of the in-school matches selected for the present study participated in other OnTrack Greenville support programs that influenced student behavior. This study originally intended to control for participation in other OnTrack Greenville support programs to address this limitation; therefore, this represents a deviation from the Sub-Grantee Evaluation Plan.

E. Next Steps

While this is a final report to satisfy Social Innovation Fund grant requirements, implementation of the program and impact evaluation will continue for two additional academic years. Given the findings presented in this report, stakeholders and researchers have several possible next steps for implementation and research.

Program stakeholders should continue to place an earlier focus on recruitment, with strong collaboration with district and feeder elementary schools. In summer 2016, BellXcel had to adjust its recruitment criteria throughout spring 2016 in order to enroll enough scholars, eventually opening up the program to all rising sixth-grade students. Recruitment efforts were more successful when preparing for summer 2017. BellXcel and their OnTrack Greenville partners should continue to monitor the enrollment targets and eligibility requirements in order to recruit and enroll the right target population of students.

Further, program leaders should continue to select enrichment courses for which there are appropriate facilities and resources. Researchers observed well-executed enrichment classes, but the cooking/nutrition courses seemed disappointing. Researchers recommend that BellXcel be more selective in considering enrichment courses and confirm that the teachers and students will have access to adequate facilities and other resources to make those courses successful.

There are additional opportunities for sharing best practices for community time and scholar dollar implementation across OnTrack Greenville sites. Community time and scholar dollars were implemented unevenly across sites, and certain sites were standouts. BellXcel should share best practices so that all schools can implement these program elements in a high quality way.

Project leaders should revisit changes to the lesson plan template and process. In summer 2017, a change in the format of lesson plans, as described by veteran staff, caused much confusion and consternation at one site. Complaints about the change and the amount of time required for staff to fill out the lesson plan template were a main theme of the teacher focus group at that site. The site's instructional coach also expressed concern. However, the format change was almost a non-issue at the other two sites. Researchers recommend that BellXcel revisit the template redesign since some teachers reported it was unnecessarily onerous. The problem may lie in how the purpose and use of the template are presented during the training, including to the instructional coaches who are tasked with supporting teachers in filling out the template and reviewing the completed lesson plans. The instructional coach at the site that reported trouble with the revised template may have over-interpreted the requirements.

In addition, researchers recommend that BellXcel revisit or better explain the pacing guides, particularly for ELA. The ELA pacing guides commonly were described as unrealistic. Some teachers believed that they were impossible to comply with and therefore did not try as hard and only considered them a minor inconvenience. Others were more frustrated. Researchers recommend revisiting the pacing guides, making some adjustments to ambitious timeframes, and including language acknowledging that teachers may need to use their professional judgment to make other adjustments. Researchers also recommend Including BellXcel teachers in the discussion of pacing guides so that more realistic timeframes are given.

Teachers also may benefit from additional support in lesson differentiation. Meeting the diverse academic needs of students is one of the most important, and most challenging, aspects of teaching. Doing so in the relatively short period of a summer program presents additional complications. Researchers recommend BellXcel continue to work towards supporting teachers in this area, including by thinking about how scholars are grouped in BellXcel classes following their initial STAR tests results or by the instructional materials provided to BellXcel teachers.

As the evaluation continues, researchers should seek a more appropriate outcome measure for course performance. Though the evaluation design will remain similar moving forward, the lack of Social Innovation Fund guidelines will give researchers more flexibility in studying the program with a less rigorous design. Modifications to the evaluation plan for summer 2019 should be a priority. These modifications should maintain as much as the existing evaluation framework as possible, but take into account the measurement issues that have challenged the study these first two years of the impact study. Evaluation efforts that contribute to the broader evidence base of summer learning also should be a priority.

In addition, researchers and project stakeholders should continue to discuss and document the implementation of school-wide policy changes at OnTrack Greenville treatment schools. As OnTrack Greenville schools continue to refine school-wide policy around attendance, behavior, and course performance, ongoing discussion and documentation of these changes will allow stakeholders and researchers to understand better and potentially tease out the impact of individual implementation partners and the impact of broader change at a systems level.

Appendix A. Study Logistics Updates

A. Institutional Review Board

There were no issues securing Institutional Review Board approval for this study. Furman University's Institutional Review Board approved and oversaw all research activities affiliated with the impact study. Furman University's IRB reviewed this research under its Expedited review process. The original application was submitted to Furman's IRB in July 2016 and approved in August 2016. Modification requests were submitted for IRB review on an ongoing basis and continuation requests were submitted annually. The school district and school personnel informed parents and guardians of the interventions and services available to their students and secured permission to provide services when necessary. Evaluators followed all parental consent and child assent protocol, as dictated by Furman University IRB guidelines and Greenville County Schools' district research protocol. These protocols detailed precisely how researchers must protect data electronically and in hard copy, and detailed informed consent procedures for both parents (parental consent) and students (child assent).

The implementation evaluation was governed by RTI International's IRB. All new RTI projects undergo an IRB needs assessment, including reviewing any specific IRB requirements for clients. Given that this was an evaluation of an educational program and students' data we de-identified, the project was exempt from a full IRB review. RTI only used aggregate data or de-identified data only in its evaluation.

The IRB evaluated the research protocol to determine whether the physical, psychological, or social risks to study participants were reasonable in relation to the anticipated benefits. After the project was approved by the RTI IRB, the IRB continued to monitor the research process to ensure that the procedures for protecting human subjects were followed. The study was reviewed by the IRB at least annually.

B. Project Timeline

There were very few modifications to the evaluation timeline for data collection, analysis, and reporting. The most notable change is that data from the South Carolina Department of Education for academic year 2017-18 were not made available in November as originally planned. At present time, researchers still have not received these data; therefore, researchers were not able to conduct statistical analyses for the state comparison group for the 2017-18 academic year for inclusion in this report.

Researchers intended to administer the OnTrack Greenville Student Pre-Survey in September of each academic year. For several reasons, pre-survey administration occurred in October instead. First, in academic year 2016-17, schools were administering the MAP assessment in September and standardized testing took precedence over data collection. In addition, researchers required the entire month of September to manage the parental consent process and prepare survey administration materials. Subsequently, researchers established a two-week survey administration window for schools in early October. In academic year 2016-17, schools were closed unexpectedly in early October due to Hurricane Michael, which delayed survey administration at some school sites. One comparison school experienced additional challenges with having adequate electronic devices for survey administration and did not complete survey administration until early November 2016.

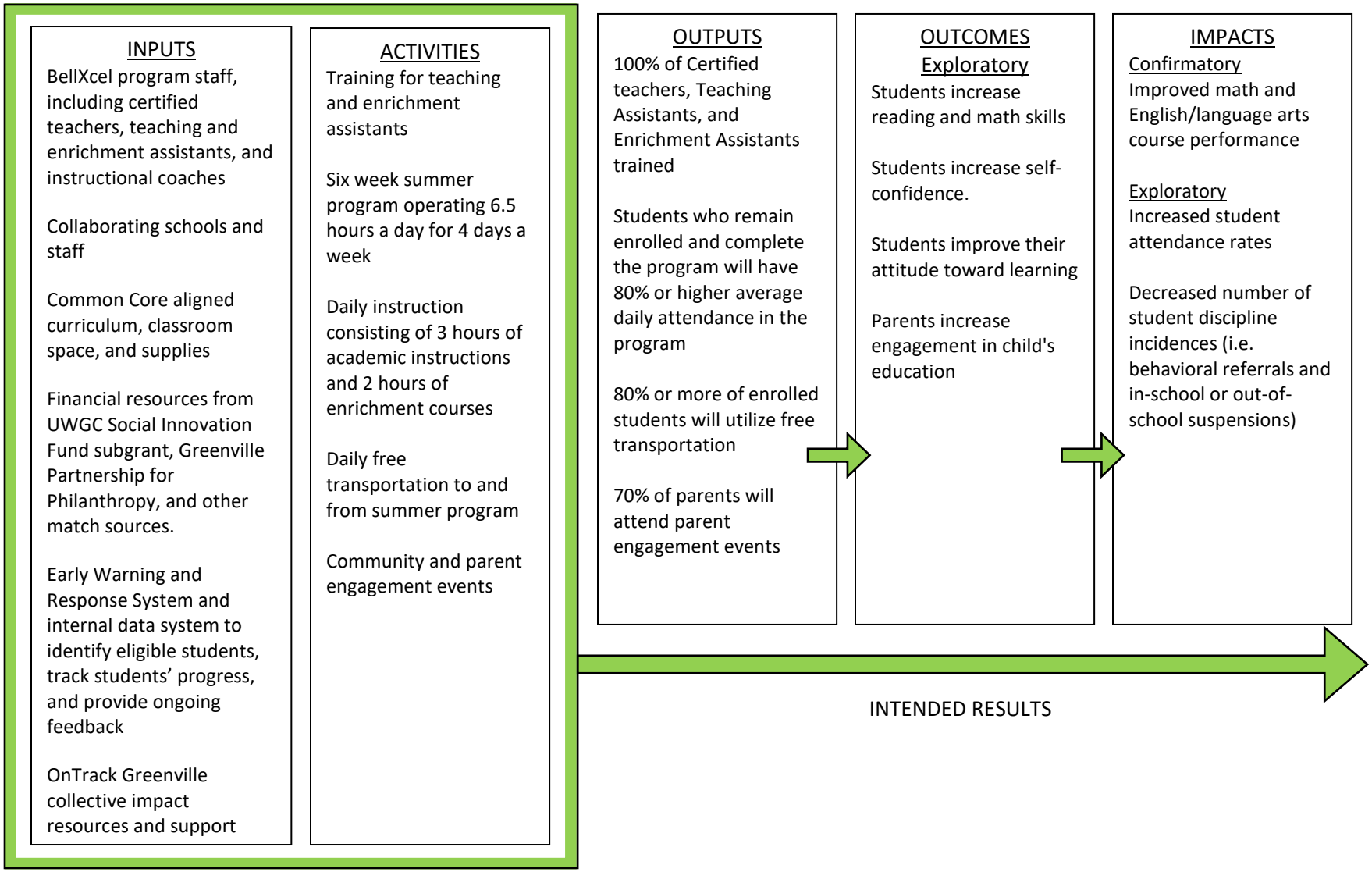
C. Project Personnel

Members of the research teams at the Riley Institute and RTI International remained constant, but there were some changes to BellXcel staff members managing the project. In January of 2017, BellXcel shifted project management from its regional Carolinas office to a dedicated Program Director for its OnTrack Greenville sites. There was some turnover among site leaders each summer, which was not unexpected and did not affect any aspects of study logistics.

D. Project Budget

Apart from the unavailability of Social Innovation Fund continuation grant monies for Year 4 and Year 5 of the project, there were no issues with or changes to the budget for this evaluation. OnTrack Greenville stakeholders have secured non-federal funding to continue the implementation and evaluation of OnTrack Greenville through academic year 2019-20.

Appendix B. Program Logic Model



Appendix C. Additional Matching Results

Appendix C provides further information on the matching process, organized by match. Each section provides: 1) additional data on the number of students matched via a participant flow chart, (2) the demographic differences between the BellXcel students who were matched and those who were not, (3) the overlap in propensities between the BellXcel and comparison students, and (4) further evidence of balance in the form of post-match standardized mean differences and variance ratios of the variables used in the matching process.

Participation Flow Chart: The BellXcel and comparison sample sizes are presented for each match. Given that each match was done independently, the number of students in each match varies, so a separate flow chart for each match is necessary. One will note that the sample sizes for the matches using survey data are much smaller than those using the administrative data. This is because researchers limited potential matches to those who responded to the student survey.

Demographic Differences between Matched and Unmatched BellXcel Students: As is evident in the participant flow charts, not all BellXcel students were matched. This has important implications for the generalizability for the results presented in this evaluation. The estimated effects of BellXcel participation are limited to those who are included in the analysis. It is possible that the effect of receiving BellXcel services is different for those who could not be matched. To get a better sense of the matching results, researchers compared the demographic characteristics of those BellXcel students who were matched to those who were not because they were missing full matching data or they did not have an available match in the given caliper. This allows one to examine how similar the BellXcel sample is to the BellXcel population on these factors.

Overlap: One goal of the matching process is for there to be substantial overlap in the propensity scores of the BellXcel students and the comparison group. To get a sense of this overlap, kernel densities were estimated for the BellXcel and comparison samples after the matching process. Frequency weights were used to account for matching with replacement. Further, the natural log of the propensity score was used in the figures, since it is not truncated at zero and one. Substantial overlap between the distributions of the BellXcel and comparison groups is evidence of good balance.

Evidence of Balance: When considering the balance of the matches, researchers considered the standardized differences between the two groups and the variance ratios. The goal was to have standardized mean differences below 0.1 and variance ratios near 1 (Steiner & Cook, 2013). If researchers found that the initial matching process created imbalanced samples, they re-estimated the propensity model using higher-order terms and interactions between the covariates (Rosenbaum & Rubin, 1984, 1985). This iterative process led to different combinations of variables being included in different matching procedures. Following Steiner and Cook (2013), the figures below demonstrate the improvement from the pre-match to the post-match balance in terms of standardized differences and variance ratios. The figures highlight that the matching process produced two very similar samples based on these factors.

Across the seven matches, the standardized mean differences for all of the variables used in the matching procedure were less than 0.1. Researchers were unable to get the variance ratios for all the

included covariates between 0.75 and 1.25 for all of the matches. However, in these cases, the variance ratios were generally only marginally beyond the thresholds (0.72, 0.69, 0.63, and 0.52).

A. AY 2016-17 Treatment School Administrative Match

Table 44. BellXcel Participant Flow Chart at Treatment Schools Summer 2016 (Administrative Data)

Study Time-point	Total number students	Number students included	Number students not included	Notes
Treatment Students				
1. Program Roster	190	---	---	
2. Appeared on School Roster	190	184	6	
3. Met Treatment Inclusion Criteria	184	104	80	Attended at least 80% of Summer Program
4. Had Full Matching Data	104	101	3	
5. Matched	101	100	1	
6. Included in Main Analyses	100	100	0	
Comparison Students				
1. School Rosters		---	---	
2. Met Treatment Inclusion Criteria	5,267	535	4732	Treatment School, Didn't participate in program, 6th grade
3. Had Full Matching Data	535	428	107	
4. Matched	428	226	202	Unique students (note: matching was done with replacement)
5. Included in Main Analyses	226	226	0	

Table 45. Which Students Got Matched? Post-Match Demographics, AY 2016-17 Treatment School Matches (Administrative Data)

	BellXcel Not Matched - Mean	BellXcel Matched - Mean	Difference	S.E. of Diff.
Black	0.41	0.31	0.10	0.07
Hispanic	0.46	0.44	0.02	0.08
White	0.11	0.18	-0.07	0.05
Other Race	0.02	0.06	-0.04	0.03
Free/Reduced Meals	0.96	0.92	0.04	0.04
Female	0.53	0.45	0.08	0.08
Special Ed.	0.26	0.25	0.01	0.07
ESL	0.39	0.36	0.03	0.07

Significance determined by chi-square tests. + 0.10 * 0.05 ** 0.01 *** 0.001

Figure 6. AY 2016-17 Treatment School Comparison Matches (Administrative Data): Overlap

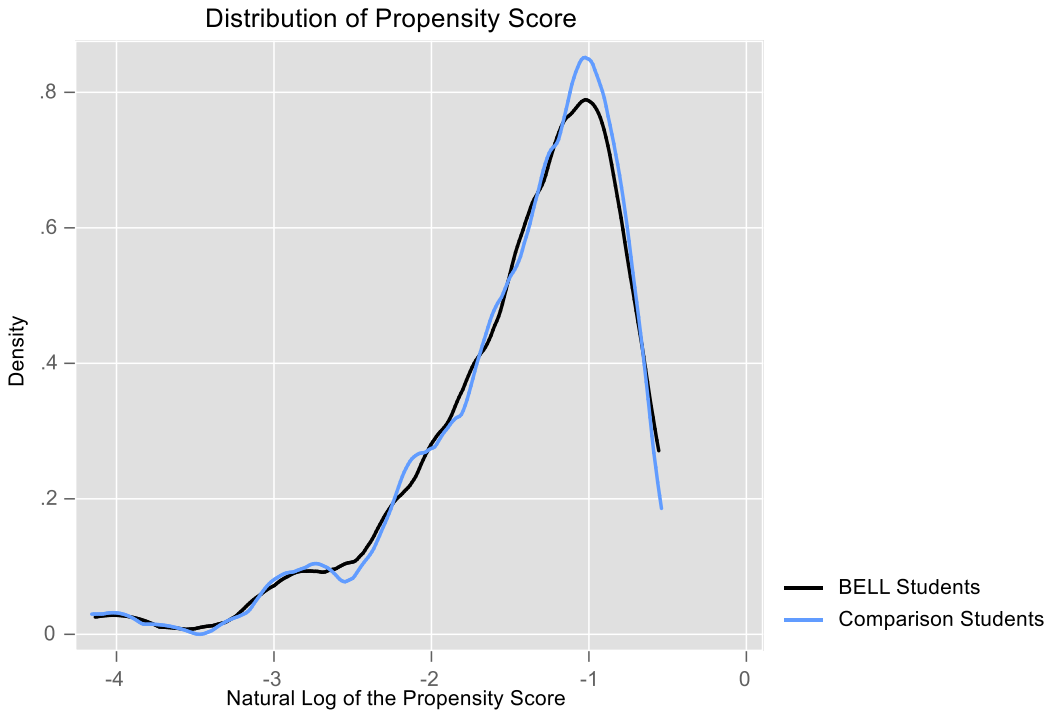
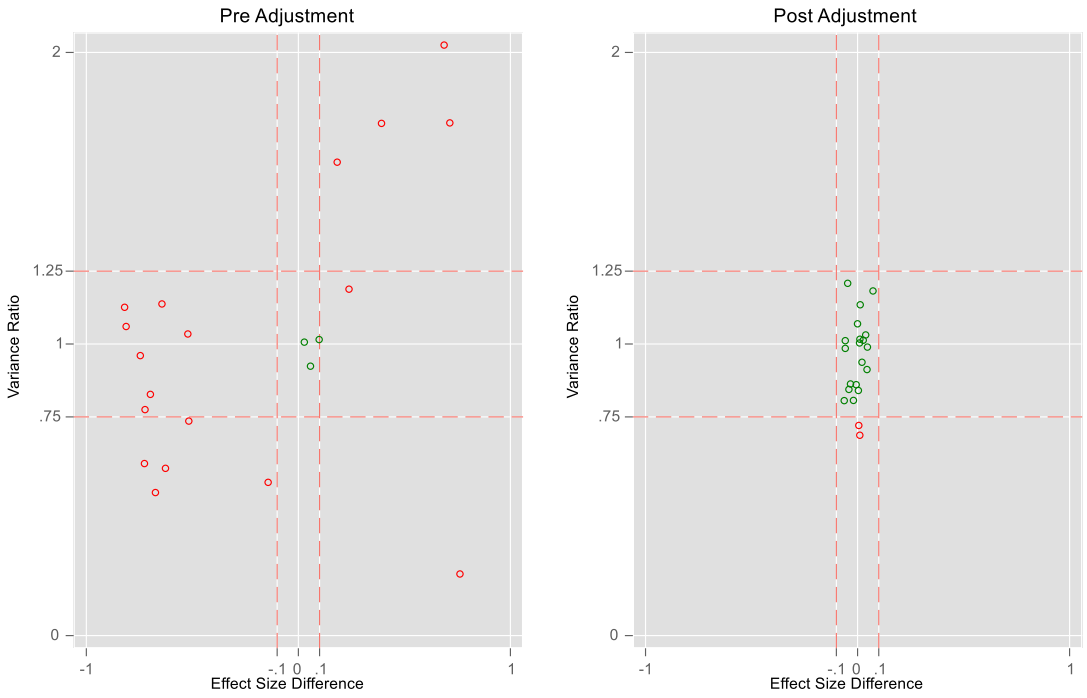


Figure 7. AY 2016-17 Treatment School Comparison Matches (Administrative Data): Standardized Differences and Variance Ratios



B. AY 2016-17 District Administrative Match

Table 46. BellXcel Participant Flow Chart at District Schools AY 2016-17 (Administrative Data)

Study Time-point	Total number students	Number students included	Number students not included	Notes
Treatment Students				
1. Program Roster	190	---	---	
2. Had Roster Data	190	184	6	
3. Met Treatment Inclusion Criteria	184	104	80	Attended at least 80% of Summer Program
4. Had Full Matching Data	104	101	3	
5. Matched	101	100	1	
6. Included in Main Analyses	100	100	0	
Comparison Students				
1. School Rosters	5,267	---	---	
2. Met Treatment Inclusion Criteria	5,267	535	4732	Treatment School, Didn't participate in program, 6th grade
3. Had Full Matching Data	535	428	107	
4. Matched	428	226	202	Unique students (note: matching was done with replacement)
5. Included in Main Analyses	226	226	0	

Table 47. Which BellXcel Students Were Matched? Post-Match Demographics, AY 2016-17 District School Matches (Administrative Data)

	BellXcel Not Matched - Mean	BellXcel Matched - Mean	Difference	S.E. of Diff.
Black	0.45	0.40	0.05	0.14
Hispanic	0.41	0.53	-0.12	0.14
White	0.12	0.07	0.05	0.09
Other Race	0.02	0.00	0.02	0.04
Free/Reduced Meals	0.95	1.00	-0.05	0.06
Female	0.52	0.60	-0.08	0.14
Special Ed.	0.26	0.27	-0.01	0.12
ESL	0.34	0.53	-0.19	0.13

Significance determined by chi-square tests. + 0.10 * 0.05 ** 0.01 *** 0.001

Figure 8. AY 2016-17 District School Comparison Matches (Administrative Data): Overlap

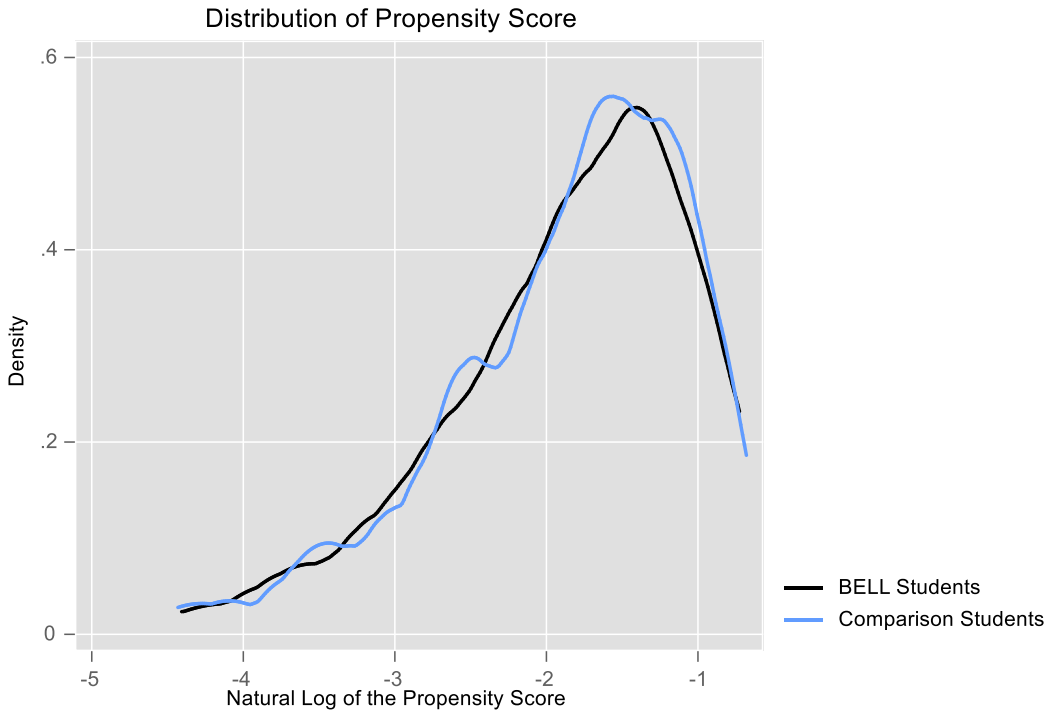
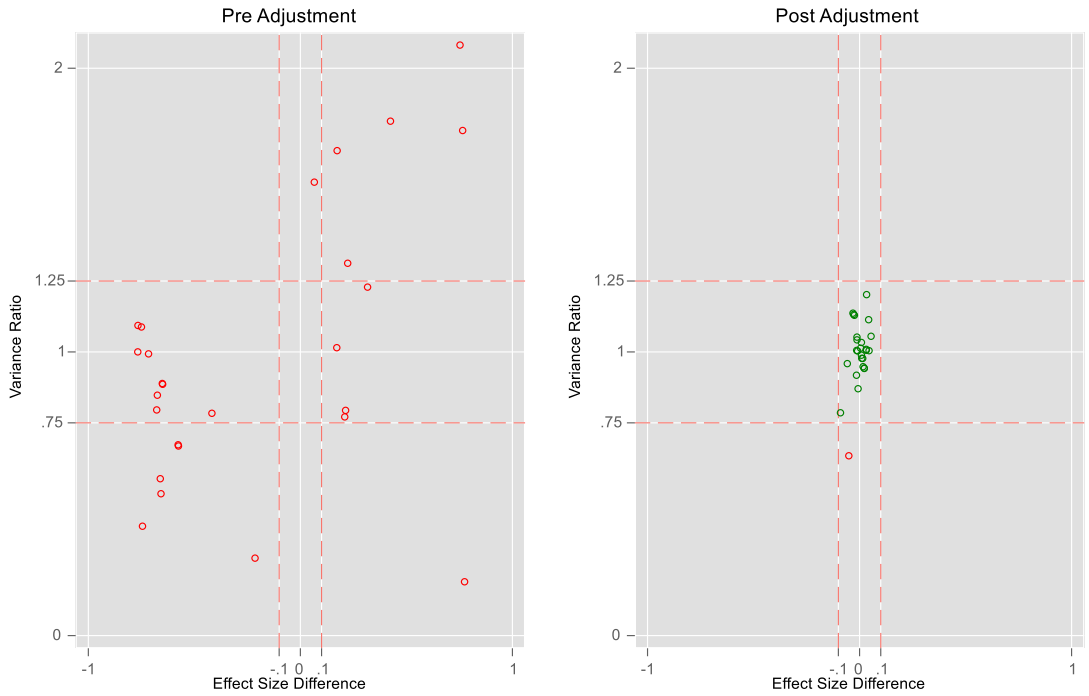


Figure 9. AY 2016-17 District School Comparison Matches (Administrative Data): Standardized Differences and Variance Ratios



C. AY 2016-17 State Administrative Match

Table 48. BellXcel Participant Flow Chart at State Schools AY 2016-17 (Administrative Data)

Study Time-point	Total number students	Number students included	Number students not included	Notes
Treatment Students				
1. Program Roster	190	---	---	
2. Had Roster Data	190	184	6	
3. Met Treatment Inclusion Criteria	184	104	80	Attended at least 80% of Summer Program
4. Had Full Matching Data	104	101	3	
5. Matched	101	98	3	
6. Included in Main Analyses	98	98	0	
Comparison Students				
1. School Rosters	797,331	---	---	
2. Met Treatment Inclusion Criteria	797,331	16,547	780,784	Title I School, Grade 6, Not in GCS district
3. Had Full Matching Data	16,547	15,056	1,491	
4. Matched	15,056	465	14,591	Unique students (note: matching was done with replacement)
5. Included in Main Analyses	465	465	0	

Table 49. Which BellXcel Students Were Matched? Post-Match Demographics, AY 2016-17 State School Matches (Administrative Data)

	BellXcel Not Matched - Mean	BellXcel Matched - Mean	Difference	S.E. of Diff.
Black	0.20	0.42	-0.22	0.23
Hispanic	0.80	0.45	0.35	0.23
White/Other	0.00	0.13	-0.13	0.15
Free/Reduced Lunch	1.00	0.90	0.10	0.14
Female	0.80	0.52	0.28	0.23
Special Ed.	0.80	0.21	**0.59****	0.19
ESL	0.80	0.35	**0.45***	0.22

Significance determined by chi-square tests. + 0.10 * 0.05 ** 0.01 *** 0.001

Figure 10. AY 2016-17 State School Comparison Matches (Administrative Data): Overlap

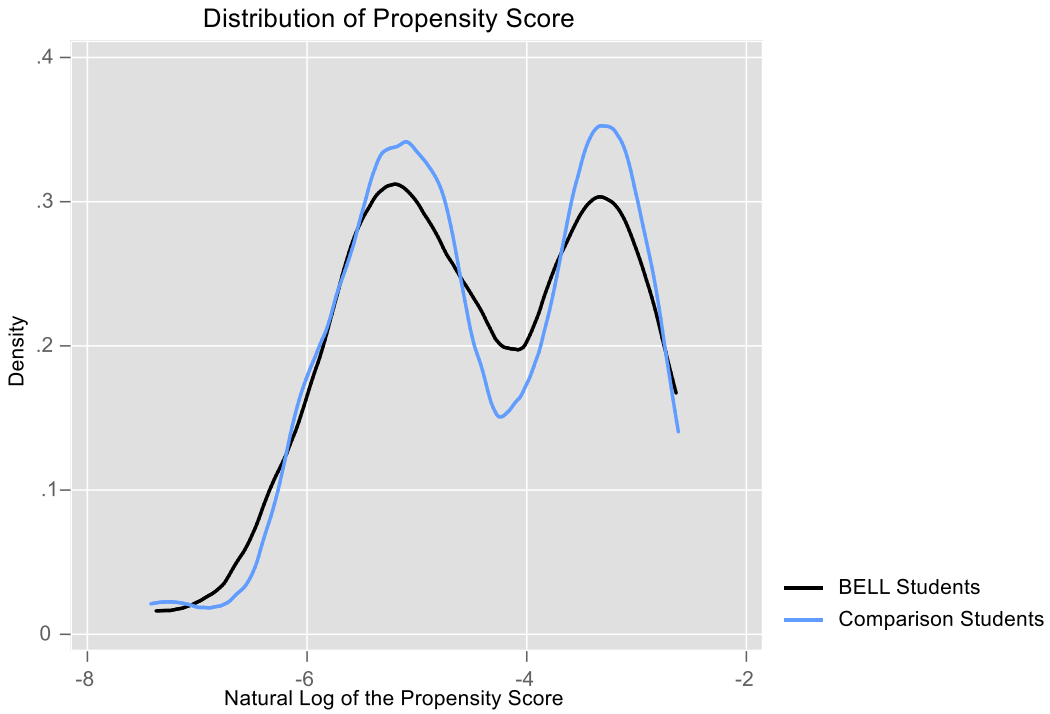
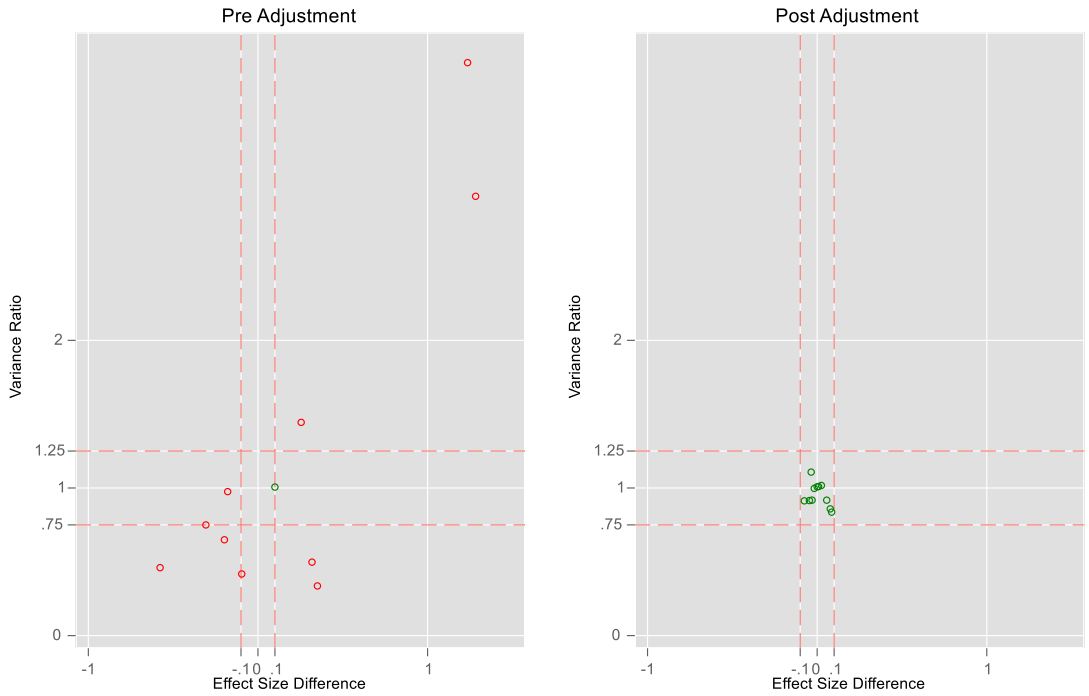


Figure 11. AY 2016-17 State School Comparison Matches (Administrative Data): Standardized Differences and Variance Ratios



D. AY 2017-18 District Administrative Match

Table 40. BellXcel Participant Flow Chart at District Schools AY 2017-18 (Administrative Data)

Study Time-point	Total number students	Number students included	Number students not included	Notes
Treatment Students				
1. Program Roster	230	---	---	
2. Had School Roster Data	230	206	24	
3. Met Treatment Inclusion Criteria	206	109	97	Attended at least 80% of Summer Program
4. Had Full Matching Data	109	106	3	
5. Matched	106	101	5	
6. Included in Main Analyses	101	98	3	
Comparison Students				
1. School Rosters	5,267	---	---	
2. Met Treatment Inclusion Criteria	5,267	1,184	4,083	District School, 6th grade
3. Had Full Matching Data	1,184	1,121	63	
4. Matched	1,121	288	833	Unique students (note: matching was done with replacement)
5. Included in Main Analyses	288	282	6	

Table 51. Which BellXcel Students Were Matched? Post-Match Demographics, AY 2017-18 District School Matches (Administrative Data)

	BellXcel Not Matched - Mean	BellXcel Matched - Mean	Difference	S.E. of Diff.
Black	0.00	0.31	-0.31	0.21
Hispanic	1.00	0.53	**0.47***	0.23
White	0.00	0.13	-0.13	0.15
Other	0.00	0.03	-0.03	0.08
Free/Reduced Meals	1.00	0.86	0.14	0.16
Female	0.40	0.51	-0.11	0.23
Special Ed.	0.60	0.18	**0.42***	0.18
ESL	0.80	0.29	**0.51***	0.05

Significance determined by chi-square tests. + 0.10 * 0.05 ** 0.01 *** 0.001

Figure 12. AY 2017-18 District School Comparison Matches (Administrative Data): Overlap

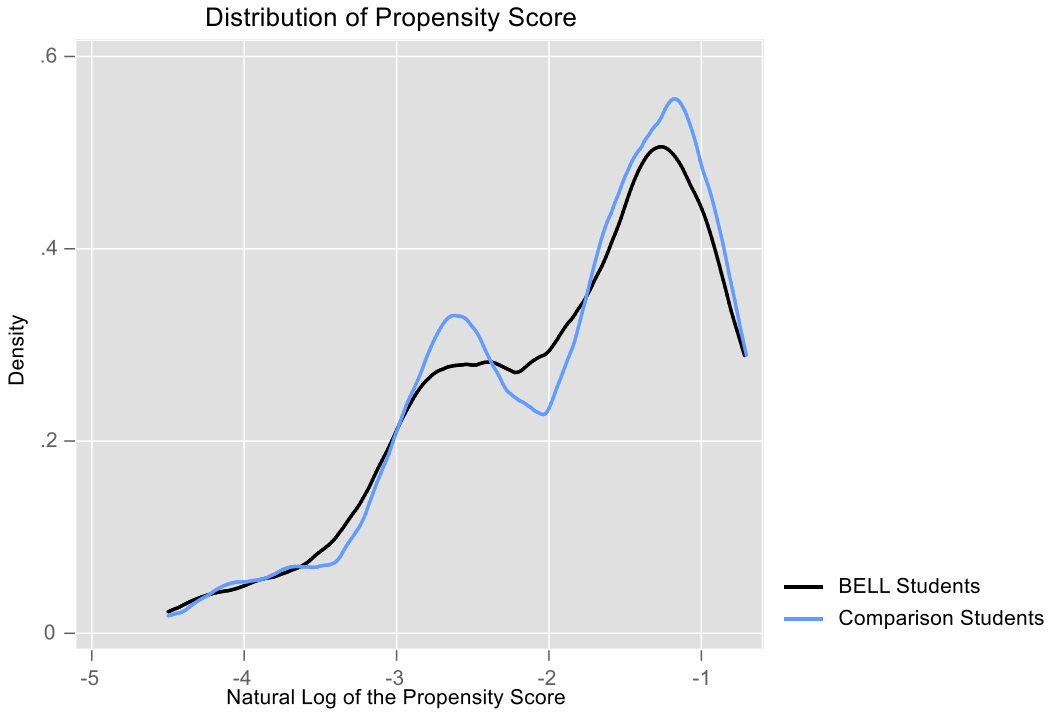
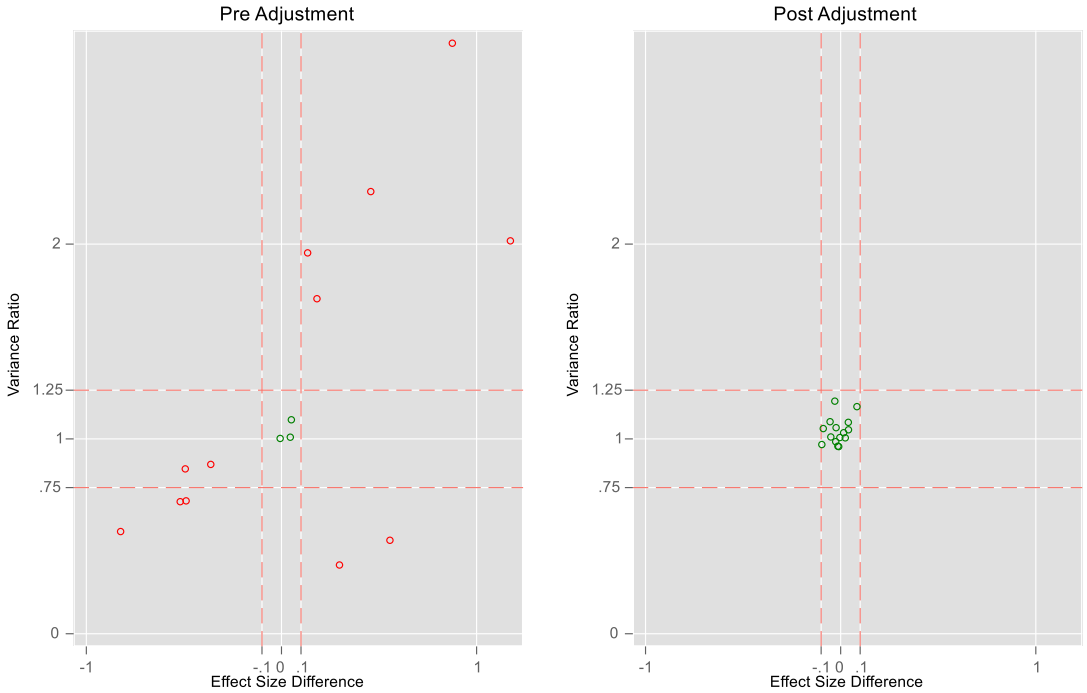


Figure 13. AY 2017-18 District School Comparison Matches (Administrative Data): Standardized Differences and Variance Ratios



E. AY 2017-18 Treatment School Survey Match

Table 52. BellXcel Participant Flow Chart at Treatment Schools AY 2017-18 (Survey Data)

Study Time-point	Total number students	Number students included	Number students not included	Notes
Treatment Students				
1. Program Roster	230	---	---	
2. Had School Roster Data	230	206	24	
3. Met Treatment Inclusion Criteria	206	109	97	Attended at least 80% of Summer Program
4. Had Full Matching Data	109	106	3	
5. Had Survey Outcomes	106	80	26	Fall 2017 survey outcomes
6. Matched	80	74	6	
7. Included in Main Analyses	74	74	0	
Comparison Students				
1. School Rosters	5,539	---	---	
2. Met Treatment Inclusion Criteria	5,539	505	5034	Treatment School, Didn't participate in program, 6th grade
3. Had Full Matching Data	505	460	45	
4. Had Survey Outcomes	460	336	124	Fall 2017 survey outcomes
5. Matched	336	175	161	Unique students (note: matching was done with replacement)
6. Included in Main Analyses	175	175	0	

Table 53. Which BellXcel Students Were Matched? Post-Match Demographics, AY 2017-18 Treatment School Matches (Survey Data)

	BellXcel Not Matched - Mean	BellXcel Matched - Mean	Difference	S.E. of Diff.
Black	0.17	0.34	-0.17	0.20
Hispanic	0.83	0.49	0.35	0.21
White	0.00	0.14	-0.14	0.14
Other	0.00	0.04	-0.04	0.08
Free/Reduced Meals	0.83	0.86	-0.03	0.15
Female	0.83	0.54	0.29	0.21
Special Ed.	0.00	0.18	-0.18	0.16
ESL	0.00	0.24	-0.24	0.18

Significance determined by chi-square tests. + 0.10 * 0.05 ** 0.01 *** 0.001

Figure 14. AY 2017-18 Treatment School Comparison Matches (Survey Data): Overlap

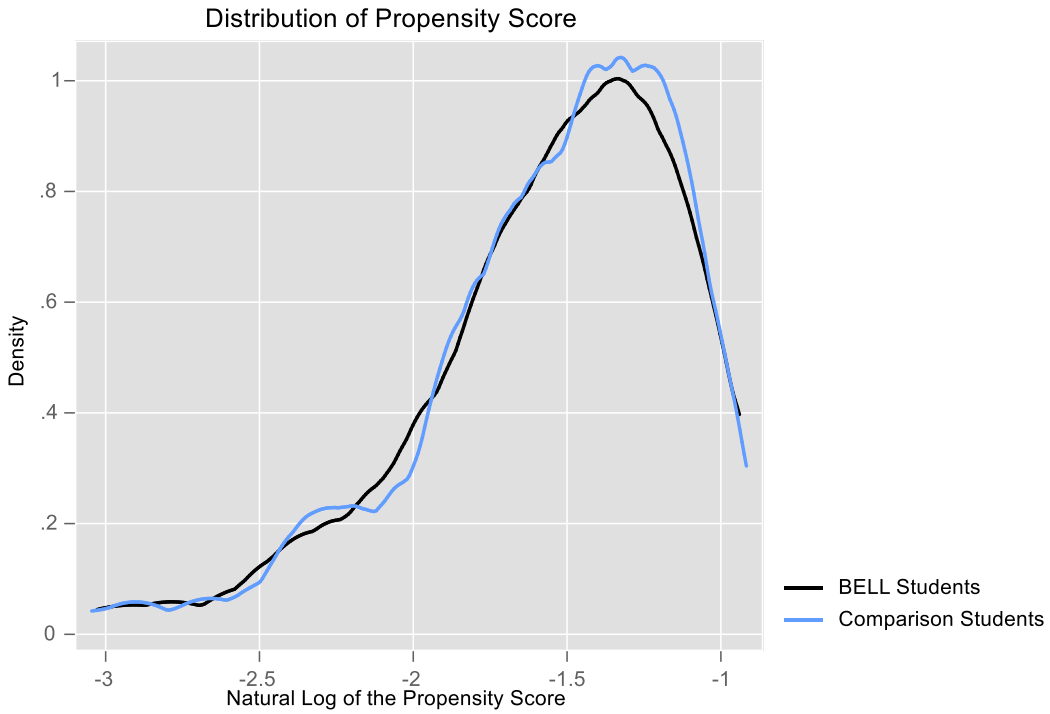
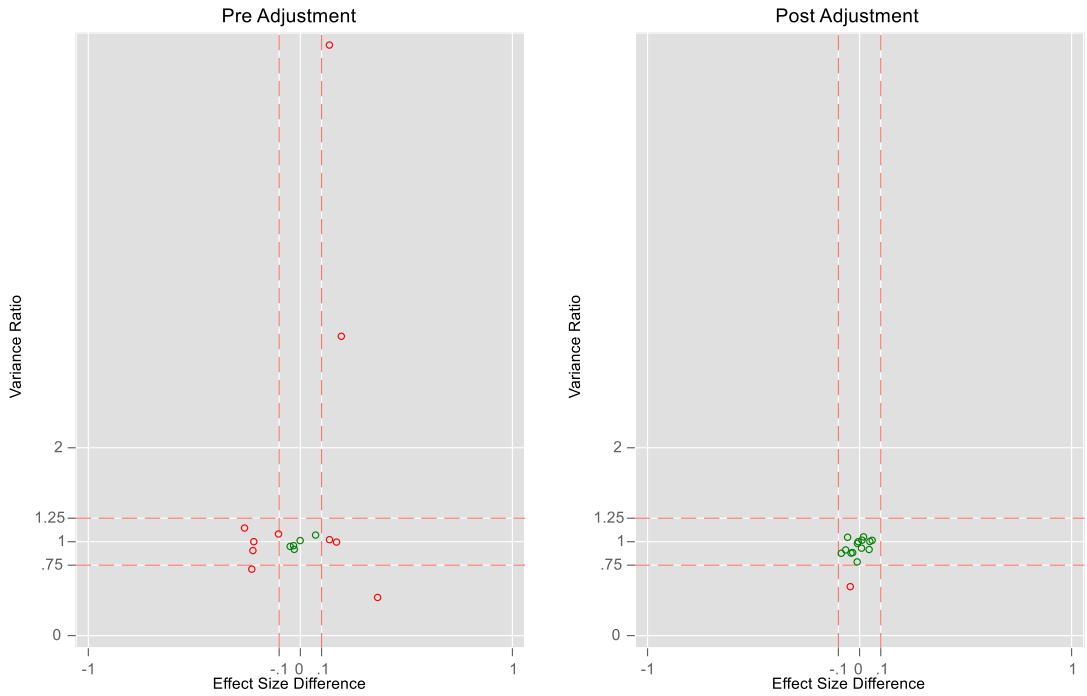


Figure 15. AY 2017-18 Treatment School Comparison Matches (Survey Data): Standardized Differences and Variance Ratios



F. AY 2017-18 District Survey Match

Table 54. BellXcel Participant Flow Chart at District Schools AY 2017-18 (Survey Data)

Study Time-point	Total number students	Number students included	Number students not included	Notes
Treatment Students				
1. Program Roster	230	---	---	
2. Had School Roster Data	230	206	24	
3. Met Treatment Inclusion Criteria	206	109	97	Attended at least 80% of Summer Program
4. Had Full Matching Data	109	106	3	
5. Had Survey Outcomes	106	80	26	Fall 2017 survey outcomes
6. Matched	80	78	2	
7. Included in Main Analyses	78	78	0	
Comparison Students				
1. School Rosters	5,267	---	---	
2. Met Treatment Inclusion Criteria	5,267	1,184	4,083	District School, 6th grade
3. Had Full Matching Data	1,184	1,121	63	
4. Had Survey Outcomes	1,121	978	143	Fall 2017 survey outcomes
5. Matched	978	250	728	Unique students (note: matching was done with replacement)
6. Included in Main Analyses	250	250	0	

Table 55. Which BellXcel Students Were Matched? Post-Match Demographics, AY 2017-18 District School Matches (Survey Data)

	BellXcel Not Matched - Mean	BellXcel Matched - Mean	Difference	S.E. of Diff.
Black	0.00	0.33	-0.33	0.34
Hispanic	1.00	0.50	0.50	0.36
White	0.00	0.13	-0.13	0.24
Other	0.00	0.04	-0.04	0.14
Free/Reduced Meals	0.50	0.87	-0.37	0.25
Female	0.00	0.58	-0.58	0.35
Special Ed.	1.00	0.14	**0.86*** **	0.25
ESL	1.00	0.21	**0.79** **	0.06

Significance determined by chi-square tests. + 0.10 * 0.05 ** 0.01 *** 0.001

Figure 16. AY 2017-18 District School Comparison Matches (Survey Data): Overlap

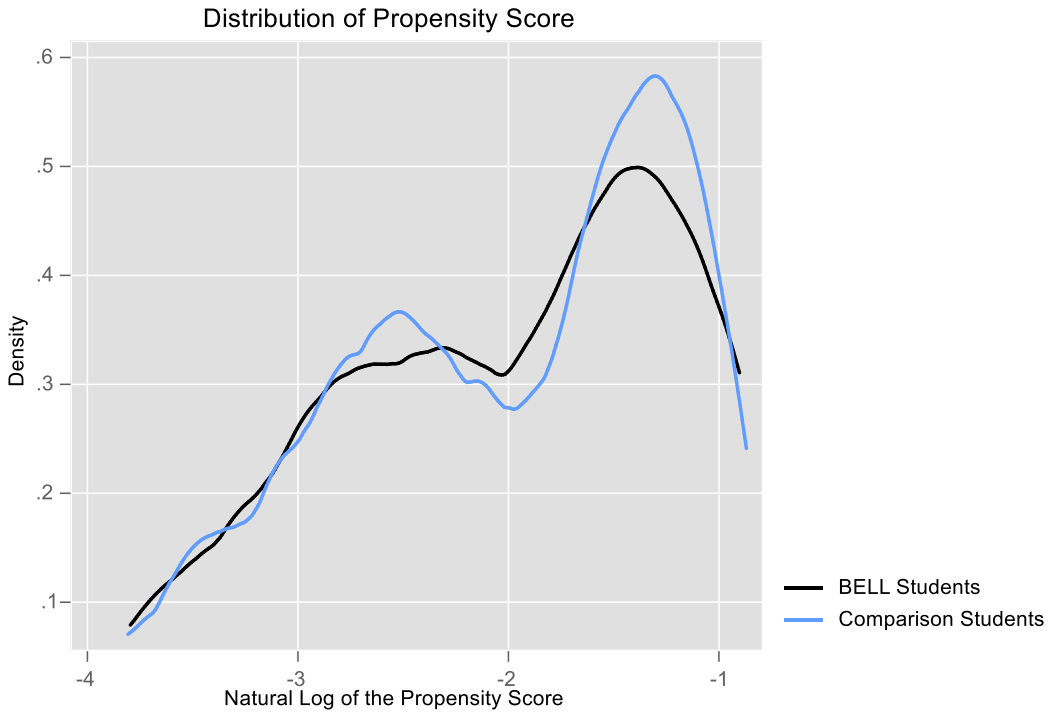
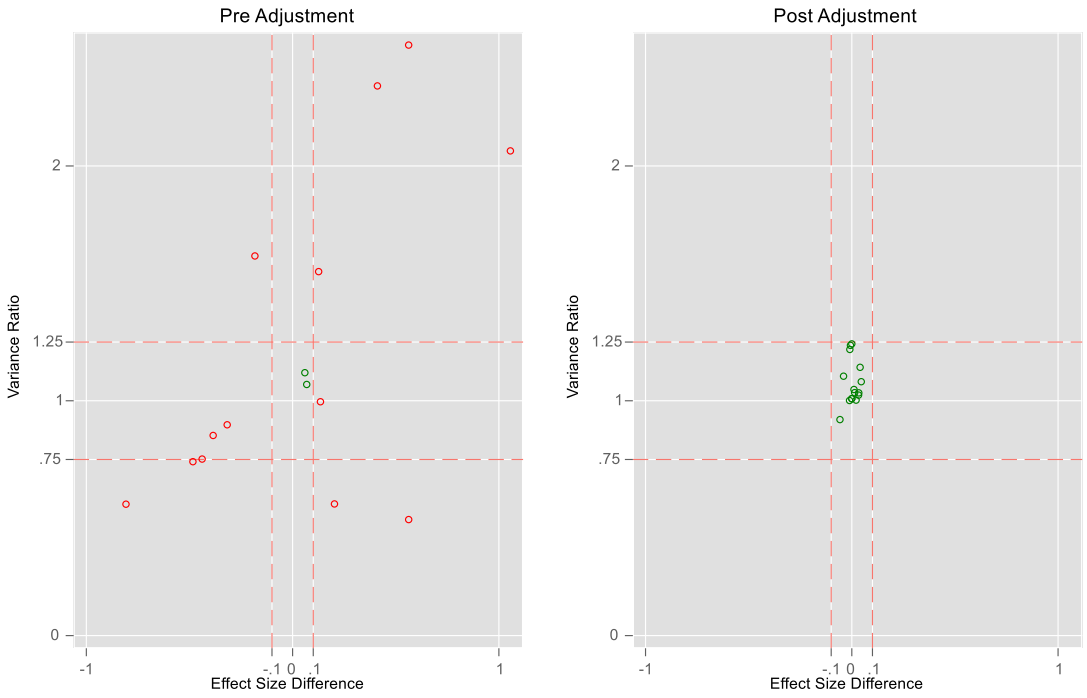


Figure 17. AY 2017-18 District School Comparison Matches (Survey Data): Standardized Differences and Variance Ratios



G. Additional Matching Results

Table 56. BellXcel Summer Program Participants vs. Student Matches AY 2017–18 (Administrative Data Match)

	BellXcel Participants (n = 100)	Student Matches: Treatment Schools (n = 500)	BellXcel Participants (n = 101)	Student Matches: District Schools (n = 505)
Black	29.0%	30.8% (-0.04)	30.7%	31.9% (-0.03)
Hispanic	55.0%	52.2% (0.06)	53.5%	53.7% (-0.00)
White/Other Race	15.0%	18.0% (-0.03)	15.9%	14.4% (0.04)
Free and Reduced Meals	87.0%	86.0% (0.03)	86.1%	86.9% (-0.02)
Female	51.0%	48.2% (0.06)	51.5%	50.3% (0.02
Special Education	16.0%	19.4% (-0.09)	17.8%	14.6% (0.08)
ESL	31.0%	29.2% (0.04)	28.7%	26.9% (0.04)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Note: Standardized mean differences are reported in parentheses. Significance tests are in comparison to the BellXcel group. Frequency weights were used to account for matching with replacement.

Table 57. BellXcel Summer Program Participants vs. Student Matches AY 2017–18 (Survey Data Match)

	BellXcel Participants (n = 74)	Student Matches: Treatment Schools (n = 370)	BellXcel Participants (n = 78)	Student Matches: District Schools (n = 390)
Black	33.8%	33.2% (0.01)	33.3%	33.9% (-0.01)
Hispanic	48.7%	45.7% (0.06)	50.0%	50.0% (0.00)
White/Other Race	17.6%	21.3% (0.09)	16.7%	16.2% (0.01)
Free and Reduced Meals	86.5%	84.9% (0.05)	87.2%	88.5% (-0.04)
Female	54.1%	51.6% (0.05)	57.7%	56.7% (0.02)
Special Education	--	--	14.1%	14.1% (0.00)
ESL	24.3%	24.6% (-0.01)	20.5%	18.7% (0.05)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Note: Standardized mean differences are reported in parentheses. Significance tests are in comparison to the BellXcel group. Frequency weights were used to account for matching with replacement.

Appendix D. OnTrack Greenville Student Survey

Please enter your survey ID number: _____

Click NEXT to continue.

You may complete the survey in English or Spanish. Please select which language you prefer. Usted puede hacer la encuesta en inglés o español. Por favor marque el idioma que prefiere.

a) English / inglés
b) Spanish / español

I am Dr. Tracy Waters from Furman University. I am conducting a study to learn about the OnTrack Greenville initiative in your school district. We are asking you to take part in the study because you are learning in a school that offers this program.

For this research, we will ask you to take a short survey. We don't think that you will encounter any problems if you participate in this survey. You can feel good about helping out with this important study. Please answer all of the questions as best you can, even if they don't seem like they apply to you.

You will not put your name anywhere on this survey. We will keep all of your answers private and will not show them to your teachers or your parents. Your answers will be stored on a password protected computer file. Only people from Furman University working on this study will see the answers students provide. When we share the results of the survey, we will never share your name or the name of your school.

You should know that:

- You do not have to be in this study if you do not want to. You won't get into any trouble with your teachers, your school, or Furman University if you say no.
- You can stop being in the study at any time.
- You can ask any questions you have, now or later. If you think of a question later, you or your parents can contact me at 864-294-3803.
- Your parents/guardians have been provided information about this study and have been given the opportunity to decline your participation.

Click NEXT to continue.

Are you willing to take this survey? By choosing "Yes," below, you acknowledge that you:

- Understand what you will be doing for this study,
- Have had all your questions answered,
- Have talked to your parent(s)/legal guardian about any questions you have about the study and,
- Agree to take part in this study.

If you choose "No," the survey will end.

a. Yes
b. No

What grade are you in?

a. 6th
b. 7th
c. 8th

[Pre-Survey] We are going to ask you about the kinds of things you did during your SUMMER BREAK, June to August 2017. Please select the answer that best reflects how often you did each of the activities listed below during the summer this year.

Responses: (1) Never; (2) Not very often (1 or 2 times); (3) Sometimes (about one a week); (4) Pretty often (a couple of times a week or more); or (5) Very often

How often…

…did you go to the library?
…did you write something like an email, letter, poem, or story?
…did you play math games or solve math problems?
…did you read a book?

[Pre-Survey] How often you did each of the activities listed below during your summer break this year?

Responses: (1) Never; (2) Not very often (1 or 2 times); (3) Sometimes (about one a week); (4) Pretty often (a couple of times a week or more); or (5) Very often

How often…

…did you play on your phone, watch TV, or play video games?
…did you do activities at a community center, YMCA, church or day camp, or Boys and Girls club?
…did you play outside?

The following statements describe your experiences at your school.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

When I study, I set goals for myself.
I keep doing schoolwork even if it is hard.
If I can't do something the first time, I keep trying until I can.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

I keep doing schoolwork even if I am bored.
When something is hard for me to do, I usually give up.
I keep doing schoolwork even if I don't like it.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

It is easy for me to get good grades in school.
I generally understand the material in my classes just as well as other students.
I am a good student.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

My homework is usually pretty easy for me.
I will be able to go as far in school as I want to go.
I can learn new things if I try.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

Doing well at school is important to me.
The things I am learning in school will be useful outside of school.
I think it is important to go to college.
I need to do well in school to accomplish my goals.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

I like school.
I participate a lot in class.
I like learning new things in school.
I feel like I matter at my school.

The following statements describe teachers at your school.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

Overall, students at my school get along well with teachers.
My teachers meet with me to talk about schoolwork and give me extra help if I need it.
My teachers really listen to what I have to say.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

Teachers at this school set a positive example for students with their actions.
My teachers notice when I am doing a good job and let me know about it.
Overall, teachers at my school try to be fair.

The following statements describe adults at your school. How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

There is at least one adult at my school who…

…really cares about me.
…tells me when I do a good job.
…notices when I am not there.

How true are the following statements?
Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

There is at least one adult at my school who…

…always wants me to do my best.
…listens to me when I have something to say.
…believes that I will be a success.

The following statements describe your experiences at your school. How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

My education will be valuable in getting the job I want.
I would be upset if I got a low grade in one of my subjects.
What I learn in school is useful for the job I want to have as an adult.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

It is important to me to get good grades.
Being a good student is important to me.
School is useful in helping me make good decisions in my life.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

I feel close to people at this school.
I am happy to be at this school.
I feel like I am a part of this school.

How true are the following statements?

Responses: (1) Not true; (2) Somewhat true; (3) Mostly true; (4) True

I feel teachers at this school treat me fairly.
I feel safe in my school.

How far *would you LIKE to go* in school with your education?

a. Some high school
b. Finish high school
c. Finish two-year/technical school
d. Finish four-year college
e. Finish graduate school or professional school
f. I don't know

How far *do you think you will actually go* in school with your education?

a. Some high school
b. Finish high school
c. Finish technical school
d. Finish college
e. Finish graduate school or professional school
f. I don't know

Physical Health refers to your diet and exercise, how often you are sick or healthy, and how your body feels.

In general, how would you describe your physical health?

(a) Excellent (b) Good (c) Fair (d) Poor

Mental Health refers to how you think and feel emotionally on a daily basis.

In general, how would you describe your emotional or mental health?

(a) Excellent (b) Good (c) Fair (d) Poor

Where do you usually go when you need to see a doctor or nurse?

a. My doctor's office
b. Emergency room
c. School health room
d. Somewhere else
e. I don't know

I am…

(a) Male (b) Female (c) Prefer not to say

I would describe myself as… (Choose all that apply)

(a) Black or African American
(b) White
(c) Spanish/Hispanic/Latino
(d) Asian or Pacific Islander
(e) American Indian or Alaska Native
(f) Other: _____
(g) Prefer not to say

--- END OF SURVEY ---

[If language = Spanish]

Yo soy la Dra. Tracy Waters de la Universidad de Furman. Yo estoy haciendo un estudio para aprender sobre el programa de OnTrack Greenville en tu distrito escolar. Te invitamos a participar en este estudio porque tú estás asistiendo a una escuela que ofrece este programa.

Para este estudio, vamos a pedir que tomes una breve encuesta. No pensamos que tendrás ningún problema si tomas esta encuesta. Tú puedes sentirse bien por ayudarnos con este estudio importante. Por favor conteste todas las preguntas lo mejor que puedas, incluso si las preguntas no te aplican.

No vamos a pedir que pongas tu nombre en ninguna parte de la encuesta. Vamos a mantener tus respuestas privadas y no vamos a compartirlas con tus maestros/as ni con tus padres. Tus respuestas serán guardadas en un archivo de computadora protegido con contraseña. Solamente personas de Furman University que trabajan en este estudio van a ver las respuestas que estudiantes proveen. Cuando compartimos los resultados de la encuesta, nunca vamos a compartir tu nombre ni el nombre de tu escuela.

Tú debes saber que:

- No tienes que participar en este estudio si no quieres. Tú no te vas a meter en problemas con tus maestros/as, tu escuela, ni la Universidad de Furman si dices “No.”
- Puedes retirar del estudio en cualquier momento.
- Puedes preguntar cualquier pregunta que tienes, ahora o después. Si piensas en una pregunta después, tú o tus padres/tutores pueden llamarme al 864-294-3803.
- Tus padres/tutores han recibido información sobre este estudio y ellos han tenido la oportunidad de declinar tu participación.

Oprima PRÓXIMO para continuar.

¿Estás dispuesto/a a tomar esta encuesta?? A escoger “Sí” debajo, tú reconoces que:

- Entiendes lo que vas a hacer en este estudio,
- Se han contestado todas tus preguntas,
- Has hablado con tus padres/tutores sobre las preguntas que tienes relacionadas con el estudio y
- Aceptas participar en este estudio.

Si escoges “No,” la encuesta va a acabar.

a. Sí
b. No

¿En qué grado estás?

a. 6º (sexto)
b. 7º (séptimo)
c. 8º (octavo)

Vamos a preguntarte sobre los tipos de cosas que hiciste durante LAS VACACIONES DE VERANO, junio a agosto 2017. Por favor marque la respuesta que mejor reflexione la frecuencia con que hiciste las siguientes actividades durante el verano de este año.

Respuestas: (a) Nunca; (b) No muy en seguido (1 o 2 veces); (c) A veces (como una vez a la semana); (d) Más o menos frecuente (un par de veces cada semana o más); (e) Muy frecuente (todos los días)

¿Con tanta frecuencia….?

…fuiste a la biblioteca?
…escribiste algo como un email, una carta, un poema o un cuento?
…jugaste juegos matemáticos o solucionar problemas matemáticas?
…leíste un libro?

¿Con tanta frecuencia hiciste cada una de las siguientes actividades durante las vacaciones de verano de este año.

Respuestas: (a) Nunca; (b) No muy en seguido (1 o 2 veces); (c) A veces (como una vez a la semana); (d) Más o menos frecuente (un par de veces cada semana o más) ; (e) Muy frecuente (todos los días)

¿Con tanta frecuencia….?

…jugaste en tu teléfono, viste televisión, o jugaste juegos videos?
…hiciste actividades en un centro comunitario, YMCA, iglesia o campo, o Club de Niños y Niñas?
…jugaste afuera?

Las siguientes frases describen tus experiencias en tu escuela.

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

¿Qué tan verdad son las siguientes frases?

Cuando estudio, yo me fijo metas para mí.
Yo sigo haciendo mi trabajo escolar incluso si es difícil.
Si no puedo hacer algo la primera vez, sigo intentando hasta que pueda.

¿Qué tan verdad son las siguientes frases

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Sigo haciendo mi trabajo escolar incluso si estoy aburrido/a.
Cuando algo es difícil para mí, normalmente me rindo.
Sigo hacienda mi trabajo escolar incluso si no me gusta.

¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Es fácil para mí sacar buenas notas en la escuela.
Generalmente entiendo el material en mis clases tan bien como otros estudiantes.
Soy buen/a estudiante.

¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Mis tareas en general son fáciles para mí.
Puedo llegar tan lejos en la escuela como quiero llegar.
Puedo aprender cosas nuevas si intento.

¿Qué tan verdad son las siguientes frases?
Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Hacer bien en la escuela es importante para mí.
Las cosas que estoy aprendiendo en la escuela serán útiles para mí afuera de escuela.
Pienso que es importante ir a la universidad.
Tengo que hacer bien en la escuela para alcanzar mis metas.

¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Me gusta la escuela.
Participo mucho en clase.
Me gusta aprender cosas nuevas en escuela.
Siento que valgo en mi escuela.

Las siguientes frases describen a los/las maestros/as en tu escuela.

¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

En general, estudiantes en mi escuela se llevan bien con maestros/as.
Mis maestros/as se reúnen conmigo para hablar de mi trabajo escolar y darme más ayuda si la necesito.
Mis maestros/as realmente escuchan lo que tengo que decir.

¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Maestros/as en esta escuela muestran un ejemplo positivo para estudiantes con sus acciones.
Mis maestros/as notan cuando estoy haciendo un buen trabajo y me dejan saber.
En general, maestros/as en mi escuela intentan ser justos/as.

Las siguientes frases describen a adultos en tu escuela. ¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Hay por lo menos un adulto en mi escuela quien…

…realmente se preocupa de mí.
…me dice cuando hago un buen trabajo.
…nota cuando no estoy presente.

¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Hay por lo menos un adulto en mi escuela quien…

…siempre quiere que yo haga mi mejor.
…escucha cuando tengo algo que decir.
…cree que voy a ser un éxito.

Las siguientes frases describen tus experiencias en tu escuela. ¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Mi educación será valiosa en obtener el trabajo que quiero.
Me pondría bravo/a si sacara una nota baja en una de mis materias.
Lo que aprendo en la escuela es útil para el trabajo que quiero tener de adulto.

¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Es importante para mí sacar buenas notas.
Ser buen/a estudiante es importante para mí.
La escuela es útil en ayudarme a hacer buenas decisiones en mi vida.

¿Qué tan verdad son las siguientes frases

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Siento cerco/a de las personas en esta escuela.
Estoy feliz estar en esta escuela.
Siento que soy parte de esta escuela.

¿Qué tan verdad son las siguientes frases?

Respuestas: (a) No verdad; (b) Un poco verdad; (c) En la mayor parte verdad; (d) Verdad

Me siento que maestros/as en esta escuela me tratan justamente.
Me siento seguro/a en mi escuela.

¿Qué tan lejos *te GUSTARIA llegar* en la escuela con tu educación?

a. Hacer una parte de la preparatorio
b. Terminar toda la preparatorio
c. Terminar la universidad de 2 años / escuela técnica
d. Terminar la universidad de 4 años / licenciatura
e. Terminar la maestría, una especialización o un doctorado
f. Yo no sé

¿Qué tan lejos *crees que actualmente vas a llegar* en la escuela con tu educación?

a. Hacer una parte de la preparatorio
b. Terminar toda la preparatorio
c. Terminar la universidad de 2 años / escuela técnica
d. Terminar la universidad de 4 años / licenciatura
e. Terminar la maestría, una especialización o un doctorado
f. Yo no sé

La salud física refiere a tu dieta y ejercicio, la frecuencia con que estás enfermo/a o saludable y como se siente tu cuerpo.

¿En general, como describes tu salud física?

(a) Excelente (b) Buena (c) Regular (d) Mala

La salud mental refiere a como piensas y te sientes emocionalmente diariamente.

¿En general, como describes tu salud emocional o mental?

(a) Excelente (b) Buena (c) Regular (d) Mala

¿Normalmente a dónde vas cuando necesitas una consulta con un/a médico/a o enfermera?

a. La oficina de mi médico/a
b. La sala de emergencias
c. La sala de salud escolar
d. Otro lado: _____
e. Yo no sé

Yo soy…

a. Varón
b. Hembra
c. Prefiero no decir

Me describo como… (Escoge todos que te describen)

a. Negro/a o Afroamericano/a
b. Caucásico/a
c. Hispano/a o Latino/a
d. Asiático/a o Isleño/a del Pacífico
e. Indio/a Americano/a o Nativo/a de Alaska
f. Otro/a: _____
g. Prefiero no decir

--- END OF SURVEY ---

Appendix E. BellXcel Scholar Survey

Thank you for coming to BELL! These are some questions about what you liked about coming to BELL. This is not a test. After you turn in this survey, no one at your site will look at it. If you have any questions, please ask your teacher.

DIRECTIONS: Please respond to each question by **checking or circling your answer choice.**

1. What grade will you be in when you go back to school in the fall? ☐ 4th ☐ 5th ☐ 6th ☐ 7th ☐ 8th ☐ 9th

2. Overall, how much did you learn about reading and English while at BELL?	*Almost Nothing*	*A little bit*	*Some*	*Quite a bit*	*A tremendous amount*
3. Overall, how much did you learn about math while at BELL?	*Almost Nothing*	*A little bit*	*Some*	*Quite a bit*	*A tremendous amount*
4. How interesting did you find your math class?	*Not at all interesting*	*Slightly interesting*	*Somewhat interesting*	*Quite interesting*	*Extremely interesting*
5. How interesting did you find your reading and English class?	*Not at all interesting*	*Slightly interesting*	*Somewhat interesting*	*Quite interesting*	*Extremely interesting*
6. How often did you participate in class activities?	*Not at all*	*A little bit*	*Some*	*Quite a bit*	*A tremendous amount*
7. Did your math skills get better or worse since coming to BELL?	*Much worse*	*Slightly worse*	*No change*	*Slightly better*	*Much better*
8. Did your reading and English skills get better or worse since coming to BELL?	*Much worse*	*Slightly worse*	*No change*	*Slightly better*	*Much better*
9. How much do you feel the adults at BELL cared about you personally?	*Almost Nothing*	*A little bit*	*Some*	*Quite a bit*	*A tremendous amount*
10. How often did you get along with your classmates while at BELL?	*Not at all*	*A little bit*	*Some*	*Quite a bit*	*A tremendous amount*
11. How often did you get along with the teachers while at BELL?	*Not at all*	*A little bit*	*Some*	*Quite a bit*	*A tremendous amount*
12. How often do students behave well in class at BELL?	*Almost never*	*Once in a while*	*Sometimes*	*Often*	*Almost always*
13. Are you more or less confident in yourself since coming to BELL?	*Much less confident*	*Slightly less confident*	*No change*	*Slightly more confident*	*Much more confident*
14. If you made a mistake or faced a challenge while at BELL, how hard did you work to overcome this mistake or challenge?	*Not hard at all*	*A little*	*Some*	*Quite a bit*	*I worked very hard*
15. How much did the staff and teachers at BELL encourage you to work hard and overcome challenges to reach your goals?	*Not at all*	*A little*	*Some*	*Quite a bit*	*A tremendous amount*
16. How often did your math teacher give you feedback that helped you learn?	*Almost never*	*Once in a while*	*Sometimes*	*Often*	*All the time*
17. How often did your reading and English teacher give you feedback that helped you learn?	*Almost never*	*Once in a while*	*Sometimes*	*Often*	*All the time*
18. Did your outlook on your future get better or worse since coming to BELL?	*Much worse*	*Slightly worse*	*No change*	*Slightly better*	*Much better*

Think about each sentence and put an X in the box that matches how you feel about it.	Yes	Maybe	No
19. I enjoyed coming to BELL	☐	☐	☐
20. I enjoyed the morning academic classes	☐	☐	☐
21. I enjoyed the afternoon enrichment classes	☐	☐	☐
22. I enjoyed the field trips	☐	☐	☐
23. I liked learning about science while at BELL? *(If you did not take science, leave this blank.)*	☐	☐	☐
24. Would you tell your friends to come to BELL?	☐	☐	☐
25. Would you come to BELL again?	☐	☐	☐

Think about each sentence and put an X in the box that matches how you feel about it.

26. While at BELL, how often did your parent(s) or guardians encourage you to read?	*Not much*	*A little bit*	*A lot*
27. While at BELL, how involved were your parent(s) or guardians in your education?	*Not much*	*A little bit more involved*	*A lot more involved*

Think about each sentence and put an X in the box that matches how you feel about it.	Agree	Disagree	Don't know
28. Intelligence is something that can change and grow in a student.	☐	☐	☐
29. Hard work is necessary to increase your knowledge and skills in a subject.	☐	☐	☐
30. You can learn from your mistakes in life.	☐	☐	☐

31. What was your favorite afternoon enrichment class at BELL? _____

32. What was your favorite field trip you went on at BELL? _____

33. Overall, what was the best part about your summer experience at BELL? _____

34. What are 1 or 2 areas where you grew or changed for the better during the summer program _____

_____

35. If you missed one or more days of the summer program, please check the reasons below. You can check more than one reason.

☐ *I had a doctor's appointment or other appointment.*	☐ *I didn't want to go to the program every day.*	☐ *I didn't like what they did at the program.*
☐ *I had trouble finding a ride to or from the program.*	☐ *I was sick.*	☐ *My family went on vacation.*
☐ *I was attending another summer program.*	☐ *Other (please specify)* _____	

36. What suggestions do you have to make the overall BELL summer experience better? _____

_____

_____

_____

Appendix F. BellXcel Staff Survey

Thank you for your service to our scholars this summer. BELL values your opinion and asks that you complete this brief survey. Your feedback will be used to evaluate the effectiveness of the program and to identify areas that need improvement. It should take 20 minutes to complete.

All surveys responses will remain confidential and will be summarized in aggregate only. Your survey responses will not affect your current or future employment with BELL. BELL Executive leaders, BELL Human Resources, and Regional and Site-Level leaders do not have access to this survey and are only provided with data in aggregate.

We will be giving away two (2) $100 Visa gift cards to randomly selected survey participants! To be eligible, you must complete the survey. Entry information is provided at the end of the survey.

If you have any questions, please contact your site Program/Site Manager. Thank you again for your service this summer and for your honest feedback. If you did not work for BELL or this summer, please do not take this survey.

Please answer the following questions about you and your experience. Remember, all responses will remain confidential.

1. Which state or program did you work in during this past BELL Summer program? [dropdown]
2. Which BELL site did you work at this summer? [dropdown: sites are ordered alphabetically by state abbreviation, then site name.]

BELL is collecting the following demographic information to better understand the background of its instructional and leadership staff.

All questions are optional and all responses are completely confidential. Opting to not respond to any question(s) will not impact your current or future employment with BELL or prevent you from filling out the rest of this survey.

3. What is your race/ethnicity? (Optional)
    a) Black/African-American
    b) White
    c) Asian
    d) Hispanic/Latino
    e) American Indian or Alaskan Native
    f) Pacific Islander or Native Hawaiian
    g) Choose not to disclose
    h) Other- please specify

4. What is your gender? (Optional)

5. What is your age? (Optional)
   a) Under 18
   b) 18 – 29 years old
   c) 30 – 49 years old
   d) 50 – 64 years old
   e) 65 years or older

6. What is the highest level of education you have attained?
   a) Some high school coursework
   b) Graduated high school
   c) Some College coursework
   d) Completed Bachelor's Degree
   e) Some Master's Degree
   f) Completed Master's Degree
   g) Some Doctorate Degree

7. Please indicate if you majored in any of the following subject areas in your education. If you have completed multiple degree programs, please pick all that apply to the various degrees listed. If you are still involved in a degree program and have <u>not</u> completed the degree, please do not pick any choice.

	Associates Degree	Bachelor's Degree	Master's Degree	Doctoral Degree
English				
Math or Statistics				
Education				
Engineering, Computer Science, Architecture				
Business Administration, Public Administration or Economics				
Science (e.g. Biology, Chemistry, Physics)				
Social science or a Humanity (e.g. history, classics, political science, sociology, anthropology)				
Psychology				
Fine Arts				
Other				

8. What grade level did you teach or oversee both during the school year? If you taught more than one grade level, please pick the HIGHEST grade level.
    a) Kindergarten
    b) 1st
    c) 2nd
    d) 3rd
    e) 4th
    f) 5th
    g) 6th
    h) 7th
    i) 8th
    j) High school grades (9th - 12th)
    k) Multiple grades (for school or program administrators)
    l) Did not teach specific grade level during this time

9. What grade level did you teach or oversee both during the BELL Summer Program? If you taught more than one grade level, please pick the HIGHEST grade level.
    a) Kindergarten
    b) 1st
    c) 2nd
    d) 3rd
    e) 4th
    f) 5th
    g) 6th
    h) 7th
    i) 8th
    j) Multiple grades (for program leaders)

10. How many years of teaching or administration experience do you have in both a traditional school year (Sept - May) setting and/or a summer program setting?

	This is my first year/ summer	1 year	2-4 years	5-9 years	10+ years teaching in this setting
Traditional School Year (September-May/June)					
Summer program					

11. Approximately how many years have you worked for the BELL Summer program?
    a) This is my first time working for the BELL Summer program
    b) 1 year
    c) 2 years
    d) 3 years
    e) 4 years
    f) 5 or more years

12. During the regular school year, did you teach or work at the same school or district where the BELL Summer program is operating? Please select the choice and statement below that best represents your experience.
    a) Yes - at the same school
    b) Yes - not at the same school, but in the same school district
    c) No - at a different urban school district
    d) No - at a different suburban school district
    e) No - at a different rural school district
    f) No - did not work for any school district during the school year

13. Are you a certified teacher or administrator?
    a) Yes
    b) No

14. When were you hired to work for the BELL Summer program?
    a) 3 or more months before program start
    b) 1-3 months before program start
    c) Less than a month before program start
    d) Less than 1 week before program start
    e) Within 1 week after program start
    f) More than 1 week after program start
    g) Don't know
    h) I am a long-term employee of BELL (>1 year of full/part-time service)

15. Which best describes your role at BELL?
    a) Program Leader (Site Manager/Program Manager/Site Coordinator/Program Director)
    b) Instructional Coach (formerly Lead Teacher)
    c) Academic Teacher - ELA
    d) Academic Teacher - MATH
    e) Academic Teacher - Math and ELA
    f) Enrichment Teacher
    g) Assistant Teacher
    h) ELL Specialist
    i) Dual Teacher (Academic and Enrichment)

<u>PROGRAM LEADERS AND INSTRUCTIONAL COACHES</u>

16. Please indicate the frequency to which you did the following activities at your site:

	Daily	1-3 times a week	1-3 times during the whole program	Rarely or never	Don't know or N/A
I informally observed classroom instructional practices.					
I formally observed classroom instructional practices using the					

Instructional Coach observation tool.					
I supported teachers' use of assessment data reports to create collaborative groups.					
I provided feedback and suggestions to teachers on instruction.					
I filled in as a substitute teacher.					
I spent time addressing administrative issues.					
I spent time addressing scholar behavioral issues.					
I spent time addressing staffing issues (conduct, tardiness, etc.)					

17. Based on the training you received from BELL, how much do you agree or disagree with the following:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Did not attend this training
Training adequately prepared me to oversee the implementation of the literacy curriculum.					
Training adequately prepared me to oversee the implementation of the math curriculum.					
Training adequately prepared me to review the lesson plans.					
Training adequately prepared me to provide guidance and coaching to elevate the performance of Teachers and Assistant Teachers.					
Training adequately prepared me to oversee test administration.					

18. Please indicate your level of agreement with the following statements with regards to the STAR Enterprise assessment system (STAR Reading, STAR Math, STAR Early Literacy):

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
The STAR Enterprise Assessment was easy to use.					
STAR Assessments were easy to					

administer to scholars.					
Scholars at my site were engaged in the STAR assessment.					
Teachers at my site used the STAR data reports to effectively differentiate instruction in the classroom.					
The STAR data reports were difficult to understand.					
The STAR data reports positively impacted classroom teaching strategies.					
Printing scholar usernames and passwords in STAR was difficult.					
I was able to add new scholars to STAR with ease.					

19. Please share any successes, challenges or other comments about your experience this summer managing a BELL Summer programsite.

ACADEMIC TEACHER

20. During the regular school year, did you have any students with English language proficiency considerations (called ELL, EL, or ELP students) in your classroom?
    a. No
    b. Yes, but just a few
    c. Yes, a majority or all of the classroom

21. During the regular school year, did you have any students with special education needs or with individualized education plans/programs (IEP's) in your classroom?
    a. No
    b. Yes, but just a few
    c. Yes, a majority or all of the classroom

22. How often did you use the following instructional strategies with scholar?

	Daily	Most days	Half the time	A few times during the program	Never
Large group instruction					
Small group instruction					
One-on-one instruction					

23. By the end of which week of the summer program did you receive STAR pre-test assessment data

reports on classroom or individual scholar performance?

a) 1st week
b) 2nd week
c) 3rd week
d) 4th week
e) 5th week
f) 6th week
g) Never received assessment report data
h) Don't know

24. Please indicate your level of agreement with the following statements with regards to the STAR Enterprise assessment system (STAR Reading, STAR Math, STAR Early Literacy):

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
STAR Assessments were easy to administer to scholars.					
Scholars at my site were engaged in the STAR assessment.					
As an academic teacher, I used the STAR data reports to effectively differentiate classroom instruction.					
The STAR data reports were easy to understand.					
The STAR data reports positively impacted classroom teaching strategies.					

25. Which of the following STAR data reports did you receive after the pre-test? (Check all that apply.)

a) Summary Report (simple table of cluster test performance)
b) Screening Report (report grouping scholars into 4 benchmark groups)
c) Instructional Planning Report - Class (groups scholars by test score and provides suggested skills)
d) Diagnostic Report (individual scholar report with diagnostic and skill information)
e) Parent Report (report specifically for parents in English and/or Spanish)
f) Other (please specify)

26. Please indicate how well you and your Assistant Teacher collaborated in the classroom and got along with each other.

	Very well	Somewhat well	Somewhat poorly	Very poorly
Collaborate around implementation of the curriculum and instructional strategies.				

Collaborate around managing scholar behavior.				
Generally get along with each other.				

CURRICULUM- LEADEARSHIP AND TEACHERS

27. Did you teach or did your site provide instruction in Math this summer?
    a) Yes
    b) No

28. Please rate your level of agreement with the following statements about the Math curriculum used at your site. If you did not teach or provide Math instruction at your site, do not answer these questions.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
The Math curriculum was appropriate based on scholars' academic needs.					
The Math curriculum design and components were easy to follow and implement.					
The Math curriculum rigorously met the expectations of the Common Core State Standards.					
The Math curriculum allowed for effective differentiation of instruction.					
The Math curriculum supported a variety of instruction; whole group, small group, and individual work.					
Teachers needed support to effectively implement the curriculum					
Supplemental materials were needed to support the curriculum.					
The Math curriculum effectively allowed teachers to mix whole group and small group instruction to keep the scholars engaged in learning.					
Scholars were engaged in the Math instruction and curriculum resources.					
The Math curriculum effectively prepared scholars for school in fall.					
The Math curriculum fostered a love of math in my scholars.					
The Math-related growth mindset activities were easy to implement					

Scholars were engaged in the Math-related growth mindset activities.					

29. Please share any suggestions on how you would improve the Math curriculum.

30. Please rate your level of agreement with the following statements about the Literacy curriculum used at your site. If you did not teach Literacy this summer, please do not answer these questions.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
The Literacy curriculum was appropriate based on scholars' academic needs.					
The Literacy curriculum rigorously met the expectations of the Common Core State Standards.					
The Literacy curriculum design and components were easy to follow and implement.					
The Literacy curriculum allowed for effective differentiation of instruction.					
The Literacy curriculum supported a variety of instruction; whole group, small group, and Teachers needed my support to effectively implement the Literacy curriculum individual work.					
Supplemental materials were needed to support the Literacy curriculum.					
The Literacy curriculum effectively allowed teachers to mix whole group and small group instruction to keep the scholars engaged in learning.					
Scholars were engaged in the Literacy instruction and curriculum resources.					
The Literacy curriculum effectively prepared scholars for school in the fall.					
The Literacy curriculum fostered a love of reading and English language arts in my scholars.					

31. Please share any suggestions on how you would improve the Literacy curriculum.

32. If you could have additional curriculum resources (for yourself or the rest of our staff), which of the following would you pick (pick your top 3)?
    a) Remedial Resources - focuses on foundational skills or lower grade-level content

b) Extension Resources - focuses on conceptual knowledge, critical thinking, or higher grade-level content
c) Hands-on & Small Group Activities
d) Project-Based Learning
e) Math Manipulatives
f) Literacy Options (eg. Leveled Library)
g) Other (please specify below)

ASSISTANT TEACHER

33. Please indicate how well you and the lead Teacher collaborated in the classroom and got along with each other.

	Very well	Somewhat well	Somewhat poorly	Very poorly
Collaborate around implementation of the curriculum and instructional strategies.				
Collaborate around managing scholar behavior.				
Generally get along with each other.				

34. Please indicate if you have any paid or volunteering (un-paid) experience in the following areas:
    a) Classroom-based academic teaching
    b) Substitute teaching
       Academic tutoring
    c) Coaching a sports team
       Working in an after school program
    d) Mentoring
       Child care

ENRICHMENT TEACHER

35. How much do you agree or disagree with the following:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
The Program Administrators supported me in the implementation of my curriculum.					

BELL clearly communicated the expectations around the enrichment classes.					

36. Which enrichment topic or curriculum did you teach this summer?

37. Based on this summer and your use of the enrichment curriculum, please rate the following items on the scale provided.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
Scholars were engaged in the curriculum					
Scholars enjoyed the activities					
Scholars became more interested in the subject matter					
The curriculum was appropriate for the grades I taught					
The training provided prepared me to deliver this curriculum					
I had the supplies I needed to do the activities					
BELL should continue to use this curriculum in the future					
Scholars were engaged in the curriculum					
Scholars enjoyed the activities					
Scholars became more interested in the subject matter					
The curriculum was appropriate for the grades I taught					
The training provided prepared me to deliver this curriculum					
I had the supplies I needed to do the activities					
I had the supplies I needed to do the activities					
BELL should continue to use this curriculum in the future					
I had the supplies I needed to do the activities					
BELL should continue to use this curriculum in the future					

PROGRAM LEADERSHIP

38. When thinking about your site, how much do you agree or disagree with the following:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
My site leaders (Program Manager and Instructional Coach) supported me in the implementation of the curriculum.					
My site leaders (Program Manager and Instructional Coach) clearly and regularly communicated the expectations around my job role and responsibilities.					
I received the tools and resources I needed from the program's leadership structure to do my job well.					
The policies for BELL staff were clearly communicated to me by my Program Manager.					
BELL's payroll process was clearly explained to me by the program administrators.					
My site leaders (Program Manager and Instructional Coach) helped me to develop my skills in managing scholar behavior.					
My site leaders (Program Manager and Instructional Coach) promoted team work at my site.					
I regularly met with my site leaders and/or other site staff to communicate site information (e.g., upcoming events, etc.)					
I regularly met with my site leaders and/or other site staff to discuss teaching, mentoring, and/or child development strategies.					
The Instructional Coach at my site gave me feedback on my instructional plans and/or delivery of instruction.					
My site leaders were highly visible around site.					

TRAINING

39. Based on the training you received from BELL, how much do you agree or disagree with the following:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
E-learning (online training) was user-friendly and structured in a way that was easy to understand.					
Training (both e-learning and classroom training) prepared me to focus on and implement growth mindset activities.					
Training prepared me to work as a collaborative team with my summer colleagues.					
Training prepared me to be a role model for my scholars.					
Training prepared me to manage behavior in my BELL classroom.					
Training inspired me to implement BELL's mission.					
BELL's online training was of a high quality.					
BELL's classroom training was of a high quality.					
Training prepared me to effectively implement the curriculum in my BELL classroom.					
Training prepared me to use assessment data to impact scholars' academic development.					

40. How often did you access and consult the following training resources during the BELL Summer program?

	Nearly every day	1 – 3 times a week	1 – 3 times during the entire program	Rarely or never	Was not aware this training was available	Don't know
BELL Teacher Handbook						
BELL University (BELL U) training modules						
BELL University library						

41. Did the BELL training provide you with all of the skills and knowledge you needed to succeed over the summer?
    a. Yes
    b. No
42. [If Q41=No] Did you access other resources to fill in any gaps in skills and/or knowledge?
    a. Yes
    b. No
43. [If Q42=Yes] What other types of resources did you access? [Check all that apply]
    a. Non-BELL textbooks or other books
    b. Resources from previous teaching jobs
    c. Help from other teachers
    d. Online resources (e.g. YouTube videos)
    e. Help from BELL instructional coaches or other BELL staff
    f. Other: _____
    g. I was not sure where to access other relevant resources

BEHAVIOR MANAGEMENT

44. Which behavior management system did you use in your classroom/site?
    a) Two-color zone system (middle school scholars)
    b) Three-color zone system (elementary school scholars) STAR system (middle school scholars)
    c) STAR system (elementary school scholars)
    d) School-based system Other (please specify)

45. Please think about the behavior management system you used in your classroom and rate the following items on the scale provided.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
The behavior management effectively managed scholar behavior in the classroom.					
Positive behavior management strategies were utilized with a focus on making good decisions.					
Recognition for good behavior and incentives are consistently administered.					
The behavior management system was consistently					
The behavior management system allowed for fair treatment of all scholars.					
The behavior management system allowed for scholars to be treated					

respectfully.					
The behavior management system allowed scholars to learn self-management.					
The BELL behavior management system is consistent with the behavior management system I use during the school year.					

46. Please indicate your choice for each of the following statements:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
I rarely have behavior problems with the scholars in this group.					
If scholars misbehave, I am comfortable dealing with it myself					
If scholars misbehave, I am comfortable calling on other BELL staff to help.					
Scholars in this class know that there will be consequences if they act out.					
I feel like I spend a lot of time trying to get scholars to settle down and stop talking.					
Most scholars in this class are good at following instructions					
This class often gets out of control.					

SCHOLAR ENGAGMENT

47. When thinking about your site, how much do you agree or disagree with the following:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
Community Time activities at my site were structured and interactive.					
Scholars participated in parent/teacher conferences.					
Scholars site participated in a meaningful community service project.					
I made at least 1 positive phone call home for every scholar in my					

cluster/site.					
We had high quality guest speakers at my site who were great role models for scholars.					
BELL reading logs were used consistently at our site to encourage reading at home.					
Scholars were engaged in my classroom.					
Community Time activities at my site were structured and interactive.					
Scholars participated in parent/teacher conferences.					

48. When thinking about your site, how much do you agree or disagree with the following:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
Students were always recognized as "scholars".					
College and career going culture was infused and present at my site.					
Opportunities for scholar "voice and choice" were provided at my site.					
Scholars were exposed to a variety of cultures, viewpoints, and perspectives.					
Scholars were provided with opportunities to work collaboratively, complete investigations and/or hands-on activities, and engage in discussions and critical thinking activities.					
Scholars enjoyed the field trips.					
Field trips, guest speakers, and cultural activities enhanced the program.					

<u>SCHOLAR ACHIEVEMENT</u>

49. Compared to the beginning of the program, how are your scholars performing in:

	Much worse	Worse	About the same	Better	Much better	N/A

Reading						
Math						
Expressing their ideas						
Having confidence in themselves						
Relating to their peers						
Focusing on their effort and hard work						
Exhibiting a positive growth mindset						

SATISFACTION

50. How much do you agree or disagree with the following:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
Based on the information I received during the recruiting and hiring process, my expectations about working for BELL were met.					
The communication, information, and support I received during the recruitment and hiring process was of high quality.					
I was hired and assigned to the exact same position that I applied for.					

51. How much do you agree or disagree with the following:

	Strongly Disagree	Disagree	Agree	Strongly Agree	Don't Know or NA
Working with the BELL Summer program met my personal and professional expectations.					
BELL program staff at my site displayed a positive attitude.					
BELL program staff at my site positively reflect BELL's mission.					
BELL program staff are diverse and					

strongly reflect the diversity of the scholars at my site.					
Working with the BELL Summer Program has helped me develop my professional skills.					
I would recommend the BELL Summer program to parents looking for an educational summer program for their children.					
I would recommend teaching in the BELL Summer program to my peers (e.g., other college students, teachers, etc.)					
I find working at BELL personally rewarding.					

52. If your circumstances stayed the same, would you want to work for the BELL Summer Program again next summer?
    a) Yes
    b) No
    c) Don't know

SATISFACTION-NEXT SUMMER

53. Please indicate any of the following reasons why you would not want to work for BELL next summer.
    a) BELL training requirements
    b) Did not meet professional expectations
    c) BELL customer service during hiring process
    d) BELL curriculum
    e) BELL customer service while employed
    f) Technology requirements of staff
    g) Site location
    h) Compensation Scholar behavior
    i) Site leadership
    j) Moving or will not be in area
    k) Do not know plans for next year yet
    l) Other (please specify)

54. Using a scale of 1 to 10, with 1 representing "Extremely Dissatisfied" and 10 representing "Extremely Satisfied," please rate your overall level of satisfaction with your teaching experience during the BELL Summer Program.

1 2 3 4 5 6 7 8 9 10

Extremely Satisfied

Extremely Dissatisfied

55. What could BELL do to most improve its summer programming for scholars?

56. What could BELL do to most improve its summer programming for staff?

57. What was your favorite or most memorable part of your summer experience with BELL?

Please provide any last thoughts or comments about your experience this summer with BELL.

Appendix G. BellXcel Parent Survey

Thank you for supporting your scholar in the BELL Summer Program. Please fill out this survey and return it to BELL.

Please answer each question about your scholar's BELL Summer experiences by putting an X in the box.

1. **What grade will your scholar be *this fall*?** ☐ K ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

2. **My scholar is a:** ☐ Male ☐ Female

3. *Since enrolling my child in BELL Summer…*	*Much worse*	*Slightly worse*	*No change*	*Slightly better*	*Much better*
a. His/her reading skills are….	☐	☐	☐	☐	☐
b. His/her math skills are…	☐	☐	☐	☐	☐
c. His/her eagerness to learn is…	☐	☐	☐	☐	☐
d. His/her self-confidence is…	☐	☐	☐	☐	☐
e. His/her attitude toward school and learning is…	☐	☐	☐	☐	☐
f. His/her ability to overcome challenges is…	☐	☐	☐	☐	☐
g. His/her eagerness to give more effort to grow and improve is…	☐	☐	☐	☐	☐

4. *Please indicate your agreement or disagreement with the statements below:*	*Agree*	*Disagree*	*Not sure*
a. Since enrolling my child in Bell, I am more involved in my child's learning.	☐	☐	☐
b. My scholar is more prepared for school in the fall as a result of BELL.	☐	☐	☐
c. My scholar enjoyed the BELL program.	☐	☐	☐
d. I would recommend the BELL Summer program to another family.	☐	☐	☐

5. *While your child has been at BELL, how often have you…*	*Not at all*	*Just a few times*	*At least once a week*	*Many times a week*	*Every day*
a. Encouraged your child to read a book	☐	☐	☐	☐	☐
b. Read a book to or with your child	☐	☐	☐	☐	☐

6. **During BELL, did staff encourage you to engage in any of the following activities?** *(Check all that apply)*

☐ Parent/teacher conference ☐ Parent orientation night ☐ Invites to field trips or speakers
☐ Take home reading activities/logs ☐ Take home math activities ☐ Resources on parent engagement
☐ Conversations (in-person, phone, email) about my child's behavior, attendance, or academics

7. **Using a scale from 1 to 10, with 1 representing "very dissatisfied" and 10 representing "very satisfied", please indicate your overall level of satisfaction with the BELL Summer program.** (*Please circle the number.*)

1 <---- ----- 2 ----- ----- 3 ----- ----- 4 ----- ----- 5 ----- ----- 6 ----- ----- 7 ----- ----- 8 ----- ----- 9 ----- -----> 10

Very dissatisfied *Very satisfied*

8. **What positive changes do you see in your child and/or family as a result of BELL? Write in the space below.**

9. **How can the BELL Summer program be improved? Write in the space below.**

DIRECTIONS: Please answer the following questions about communication with BELL staff and about BELL parent events by **putting an X in the box with your answer**

	Never	*A few times*	*About once a week*	*More than once a week*
10. How often did your child's BELL teacher(s) communicate with you about your child throughout the summer?	☐	☐	☐	☐
11. How often did other BELL staff (e.g. Program Manager) communicate with you about your child throughout the summer?	☐	☐	☐	☐

12. Overall, how satisfied were you with the amount and quality of communication with the BELL teachers and staff?

1 <---- ----- 2 ----- ----- 3 ----- ----- 4 ----- ----- 5 ----- ----- 6 ----- ----- 7 ----- ----- 8 ----- ----- 9 ----- -----> 10

Very dissatisfied *Very satisfied*

13. If you attended the parent orientation, please choose whether you agree or disagree with the following statements:

The BELL parent orientation …	*Agree*	*Disagree*	*Not sure*
a. Helped me to understand my scholar's experience in the BELL program.	☐	☐	☐
b. Was well planned and had a clear agenda.	☐	☐	☐
c. Provided the opportunity for me to connect with BELL teachers and staff.	☐	☐	☐
d. I did not attend the BELL parent orientation	☐	☐	☐

14. If you did <u>not</u> attend the parent orientation, please indicate why you did not attend. (Check all that apply.)

☐ Another family member attended ☐ Schedule conflict
☐ I did not think it would be useful ☐ I was not aware of the event
☐ Trouble finding transportation or childcare ☐ Other _____

15. If you attended the family week event, please choose whether you agree or disagree with the following statements:

The BELL family week event…	*Agree*	*Disagree*	*Not sure*
a. Helped me to understand my scholar's experience in the BELL program.	☐	☐	☐
b. Was well planned and had a clear agenda.	☐	☐	☐
c. Provided the opportunity for me to connect with BELL teachers and staff.	☐	☐	☐
d. I did not attend the family week event.	☐	☐	☐

16. If you did not attend the family week event, please indicate why you did not attend. (Check all that apply.)

☐ Another family member attended ☐ Schedule conflict
☐ I did not think it would be useful ☐ I was not aware of the event
☐ Trouble finding transportation or childcare ☐ Other_____

17. What suggestions do you have for improving future parent orientations and/or family week events? Write below.

18. If your scholar never rode the bus provided by BELL or only rode it occasionally, what were the reasons? (Check all that apply.)

☐ I did not know a bus was available ☐ My scholar walked to the summer program
☐ A family member or other adult drove my scholar to the summer program ☐ The bus times and/or stop locations were not convenient
☐ My scholar often rode the bus ☐ Other _____

19. How did you hear about the BELL Program? *(Phone, Letter, Text, etc.)* _____

20. What was the biggest factor in choosing BELL for your scholar(s)? _____

Appendix H. Program Manager Interview Protocol

1. What is your role at BELL?
    a. Has your role changed over time?

2. How does BELL implement and manage all of the components of the summer program?
    a. Please discuss roles and responsibilities of key staff for each of the following program components:
        - Community Time
        - Classroom set up/management
        - Behavioral management
        - Teaching and learning: Academics
        - Teaching and learning: Enrichment
        - Parent engagement

3. How do you monitor what's going on in the program?

4. What types of enrichment courses are offered?
    a. How were the enrichment courses selected?
    b. How were teachers recruited for teaching enrichment courses?
        i. Recruited to teach specific enrichment classes or recruited to teach one of multiple options?

5. What is the teacher/TA arrangement for academic and enrichment instruction?
    a. Do TA's stay with the students or teachers during the academic classes?
    b. Do TAs stay with the students or teachers during the enrichment classes?

6. Please describe BELL's attendance policy is being implemented in this district.
    a. How is this policy enforced at the site level?
    b. How do is this policy communicated to parents and students?

7. Please describe the behavior management policy being implemented in this district.
    a. How is this policy enforced at the site level?
    b. To what extent is there variation in implementation across teachers/TAs in using the behavioral model?
    c. What types of disciplinary problems does BELL face?

8. How does BELL address these problems?
    a. What is your assessment of the BELL behavioral model? (Probe: age appropriateness/effectiveness)

9. How do the scholars experience the broader community through the BELL summer program?
    a. What was your level of involvement with the local community prior to BELL?

b. Did you already have existing community ties or do you have to develop them after taking this position at BELL?

10. How does the BELL program engage parents?

11. How has each program component been implemented so far?
    a. How does BELL evaluate program quality and fidelity to the model?
    b. How is the data BELL gathers on program quality used?
    c. How has this district's TA arrangement been working so far?
        i. Is it being implemented consistently across clusters and grades?
    d. Would you change it going forward?
    e. How has the academic instruction been going?
        i. Is it being implemented consistently across clusters and grades?
        ii. Would you change it going forward?
    f. How has the enrichment implementation been going?

12. Is it being implemented consistently across clusters and grades?
    a. Would you change it going forward?
    b. How has community time/community engagement been going?

13. Is it being implemented consistently across clusters and grades?
    a. Would you change it going forward?

14. What is your assessment of your relationship with your staff?

15. What advice would you give to BELL in terms of the overall program model?

Appendix I. Teacher Focus Group Protocol

1. Why did you decide to become a BELL summer program teacher?

2. What type of preparation did you receive for teaching in the BELL summer program?

3. How would you assess the training provided by BELL for teachers?
    a. Usefulness, scope, length?
    b. Things you would not change?
    c. Ideas for improvement?

4. How would you assess the reading component of the BELL program?
    a. Quality/appropriateness of resources
    b. Length of instructional time

5. Do you make adaptations to the curricula?

6. How do you typically use your teaching assistants in the reading class?

7. Have you noticed any patterns in terms of reading learning by particular demographic groups of students (e.g., gender, race, grade level)?

8. How would you assess the writing component of the BELL program?
    a. Quality/appropriateness of resources
    b. Length of instructional time

9. Do you make adaptations to the curricula?
    a. If so, why?

10. How do you typically use your teaching assistants in the writing class?

11. Have you noticed any patterns in terms of writing learning by particular demographic groups of students (e.g., gender, race, grade level)?

12. How would you assess the math component of the BELL program?
    a. Quality/appropriateness of resources?
    b. Length of instructional time?

13. Do you make adaptations to the curricula?
    a. Why?

14. How do you typically use your teaching assistants in math instruction?

15. Have you noticed any patterns in terms of math learning by particular groups of students (e.g., gender, race, grade level)?

16. What types of enrichment classes do you teach/lead? How were these social enrichment activities selected?

17. How would you assess student engagement in enrichment?
    a. What important factors influence student engagement?
        i. Self-selection of courses?
        ii. Parental support?
        iii. Individual motivation?
        iv. Teacher investment?

18. How were students placed into enrichment courses?
    a. What is your assessment of this course placement/selection process?

19. How do you typically use your teaching assistants in this area?

20. Have you noticed any learning patterns by particular groups (e.g., gender, race, grade level)?

21. What is your assessment of the BELL's efforts to engage parents?
    a. How does it compare to other models of parental engagement that are familiar to you?

22. How would you describe the management style of your site director?
    a. One word or short descriptive phrase?
    b. Do you feel this style is generally effective with the teachers and TAs at this site?

23. What is your relationship with the TAs?
    a. How do you typically utilize their assistance?
    b. How do you typically communicate with them?
    c. Do you have sufficient interaction/planning time with TAs?
        i. At some BELL sites, TAs remain with the same students all day. At other sites, the TA remains with the same teacher all day. What is your assessment of each approach?

24. How would you describe your relationships with your students?
    a. Do you have a good sense of the students?
        i. Feel like you receive enough background information to effectively educate them?

25. How would you assess student engagement, overall?
    a. What key factors influence student engagement?
        i. Self-selection of courses?
        ii. Mandatory status?
        iii. Parental support?

iv. Individual motivation?
v. Teacher investment?

26. Please describe the behavior management policy being implemented in this district.
    a. How is this policy enforced at your BELL site?

27. To what extent is there variation in implementation across teachers/TAs in using the behavioral model

28. What are the most prevalent disciplinary problems that you face?
    a. How do you address these problems?
    b. What is your assessment of the BELL behavioral model?
        i. Age appropriateness?
        ii. Effectiveness?

29. Do you feel you are provided with the resources you need from program administrators to be a successful instructor in the BELL summer program?

30. If not, what additional resources or support do you need?

31. What are your biggest challenges as a BELL summer program instructor?

32. What is your overall assessment of the BELL programmatic structure and philosophy? (Probe: congruence/disconnect between concept and implementation)

33. Reflecting on your experiences so far, what suggestions would you offer to BELL summer program administrators in terms of program design and implementation?

Appendix J. Unmatched Regressions

Appendix J presents the coefficients, standard errors, and effect sizes for the BellXcel treatment variable from the unmatched regressions. Ordinary least square and logistic regressions were estimated using all observations for which complete data were available. The tables below are equivalent to the tables in Section V.D; however, these regressions include the full sample, rather than just the matched students. In the unmatched regressions, researchers controlled for race, gender, grade level, English proficiency, special education status, free and reduced meal eligibility, and baseline outcome variables when available. The covariate-adjusted BellXcel regression coefficient and the robust standard errors are presented in the tables. The coefficients were converted into Cohen's d, a measure of effect size. By comparing the unmatched results to the matched comparisons, one can see how the matching process changed the results.

Researchers compared course performance outcomes between BellXcel students and students attending treatment, district, and state schools in the 2016-17 academic year. As shown in Table 58, when compared to students attending treatment schools, BellXcel scholars had significantly lower scores on the MAP Reading ($p < 0.05$)) assessment. When compared to students attending district schools, BellXcel scholars received significantly lower scores on the MAP Math ($p < 0.001$) and MAP Reading ($p < 0.001$) assessments.

Table 58. Confirmatory Impact Unmatched Regression Results for Course Performance AY 2016-17

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools		BellXcel vs. Comparison Students: State	
	Unmatched Regression (n = 599)	Effect Size	Unmatched Regression (n = 1,221)	Effect Size	Unmatched Regression (n = 15,725)	Effect Size
MAP RIT - Math Fall 2016	-1.97 (1.41)	-0.12	**-9.61*** ** (1.33)	-0.58	---	---
MAP RIT - Reading Fall 2016	**-4.13*** (1.87)	-0.25	**-12.0*** ** (1.79)	-0.78	---	---

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; BellXcel vs. Treatment N = 600; BellXcel vs. District N = 564; BellXcel vs. State N = 588

Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared attendance outcomes between BellXcel students and students attending treatment, district, and state comparison schools in the 2016-17 academic year. As shown in Table 59, when compared to students attending treatment schools, BellXcel students had significantly lower rates of chronic absenteeism ($p < 0.05$) and significantly higher rates of average daily attendance ($p < 0.001$) in the 2016-17 academic year. When compared to students attending district schools, BellXcel scholars had higher rates of average daily attendance ($p < 0.10$). additionally, when compared to students attending state schools, BellXcel scholars had higher rates of average daily attendance ($p < 0.05$) and lower rates of chronic absenteeism ($p < 0.10$).

Table 59. Exploratory Impact Unmatched Regression Results for Attendance AY 2016-17

	BellXcel vs. Comparison Students - Treatment Schools		BellXcel vs. Comparison Students – District Schools		BellXcel vs. Comparison Students – State Schools	
	Unmatched Regression (n = 638)	Effect Size	Unmatched Regression (n = 1,261)	Effect Size	Unmatched Regression (n = 16,651)	Effect Size
Average Daily Attendance	**1.45*** (0.42)	0.29	**0.73†** (0.38)	0.16	**0.87*** (0.36)	0.16
Chronically Absent	**-1.01*** (0.41)	-0.55	-0.57 (0.40)	-0.31	**-0.74†** (0.39)	-0.40

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ BellXcel vs. Treatment School N= 588; BellXcel vs. District N= 564; BellXcel vs. State N = 588

Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared attendance outcomes between BellXcel students and students attending treatment and district comparison schools in the 2017-18 academic year. As shown in Table 60, when compared to students attending treatment schools, BellXcel students had significantly lower rates of chronic absenteeism ($p < 0.10$) and significantly higher rates of average daily attendance ($p < 0.001$) in the 2017-18 academic year. When compared to students attending district schools, BellXcel scholars had higher rates of average daily attendance ($p < 0.001$).

Table 60. Exploratory Impact Unmatched Regression Results for Attendance AY 2017-18

	BellXcel vs. Comparison Students - Treatment Schools		BellXcel vs. Comparison Students – District Schools	
	Unmatched Regression (n = 611)	Effect Size	Unmatched Regression (n = 1,290)	Effect Size
Average Daily Attendance	**1.74*** (0.34)	0.41	**1.53*** (0.30)	0.35
Chronically Absent	**-0.82†** (0.48)	-0.45	-0.61 (0.47)	-0.34

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; BellXcel vs. Treatment N = 600; BellXcel vs. District N = 606

Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared behavioral outcomes between BellXcel scholars and students attending treatment, district, and state schools during the 2016-17 academic year. When compared with students attending treatment schools, BellXcel scholars received fewer overall days of OSS during the 2016-17 academic year ($p < 0.10$). When compared with students attending district schools, BellXcel scholars were more likely to receive at least one behavioral referral in the 2016-17 academic year ($p < 0.10$), but they received fewer hours of ISS ($p < 0.01$) and fewer days of OSS ($p < 0.05$) in this same time period. Additionally, when compared to students attending state comparison schools, BellXcel students received fewer hours of ISS in the 2016-17 academic year ($p < 0.05$).

Table 61. Exploratory Impact Unmatched Regression Results for Behavior AY 2016-17

	BellXcel vs. Comparison Students - Treatment Schools		BellXcel vs. Comparison Students – District Schools		BellXcel vs. Comparison Students – State Schools	
	Unmatched Regression (n = 639)	Effect Size	Unmatched Regression (n = 1,266)	Effect Size	Unmatched Regression (n = 16,651)	Effect Size
Any Behavioral Referral	0.17 (0.22)	0.10	**0.38†** (0.21)	0.21	0.17 (0.20)	0.10
# Behavioral Referrals	-0.27 (0.30)	-0.08	0.07 (0.26)	0.02	-0.26 (0.26)	-0.06
Any ISS	0.06 (0.28)	0.03	-0.29 (0.26)	-0.16	-0.14 (0.25)	-0.08
# Hours ISS	-0.53 (0.70)	-0.07	**-2.47**** (0.69)	-0.23	**-0.19*** (0.10)	-0.12
Any OSS	-0.01 (0.26)	-0.01	0.35 (0.25)	0.19	0.13 (0.24)	0.07
# Days OSS	**-0.36†** (0.20)	-0.14	**-0.45*** (0.19)	-0.14	-0.07 (0.09)	-0.05

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ BellXcel vs. Treatment School N= 588; BellXcel vs. District N= 564; BellXcel vs. State N = 588

Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Table 62. Exploratory Impact Unmatched Regression Results for Behavior AY 2017-18

	BellXcel vs. Comparison Students - Treatment Schools		BellXcel vs. Comparison Students – District Schools	
	Unmatched Regression (n = 614)	Effect Size	Unmatched Regression (n = 1,293)	Effect Size
Any Behavioral Referral	-0.13 (0.21)	-0.07	0.18 (0.20)	0.10
# Behavioral Referrals	-0.34 (0.33)	-0.09	-0.02 (0.31)	-0.01
Any ISS	0.01 (0.26)	0.01	-0.15 (0.25)	-0.08
# Hours ISS	0.20 (0.72)	0.03	**-2.35**** (0.77)	-0.18
Any OSS	**-0.57*** (0.27)	-0.31	0.08 (0.26)	0.04
# Days OSS	**-0.59*** (0.24)	-0.20	-0.32 (0.23)	-0.09

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; BellXcel vs. Treatment N = 600; BellXcel vs. District N = 606

Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared behavioral outcomes between BellXcel scholars and students attending treatment and district schools during the 2017-18 academic year. When compared with students attending treatment schools, BellXcel scholars were less likely to receive any OSS ($p < 0.05$), and they received fewer overall days of OSS ($p < 0.05$) during the 2017-18 academic year. When compared with students attending district schools, BellXcel scholars received fewer hours of ISS ($p < 0.01$) in this same time period.

Researchers compared outcomes on student self-confidence between BellXcel scholars and students attending treatment and district schools during the 2016-17 academic year. As shown in Table 63, there were no differences in student self-confidence between BellXcel scholars and students attending treatment schools in the 2016-17 academic year. However, when compared to students attending district schools, BellXcel scholars reported lower levels of academic self-confidence ($p < 0.10$).

Table 63. Exploratory Outcome Unmatched Regression Results for Student Self-Confidence Fall 2016

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched Regression (n = 452)	Effect Size	Unmatched Regression (n = 912)	Effect Size
Academic Perseverance	-0.06 (0.08)	-0.10	-0.06 (0.08)	-0.10
Academic Self-Confidence	-0.02 (0.07)	-0.04	**-0.12†** (0.07)	-0.22

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 390.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Table 64. Exploratory Outcome Unmatched Regression Results for Student Self-Confidence Fall 2017

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched – Regression (n = 445)	Effect Size	Unmatched – Regression (n = 1,108)	Effect Size
Academic Perseverance	**-0.15†** (0.08)	-0.25	**-0.16*** (0.07)	-0.27
Academic Self-Confidence	**-0.21**** (0.08)	-0.38	**-0.26***** (0.07)	-0.47

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses. . BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared outcomes on student self-confidence between BellXcel scholars and students attending treatment and district schools during the 2017-18 academic year. As shown in Table 64, when compared to students attending treatment schools, BellXcel scholars had significantly lower levels of academic perseverance ($p < 0.10$) and academic self-confidence ($p < 0.001$) during the 2017-18 academic year. additionally, when compared to students attending district schools, BellXcel scholars

reported lower levels of academic perseverance ($p < 0.05$) and academic self-confidence ($p < 0.001$) during the same time period.

Researchers compared student attitude toward learning between BellXcel scholars and students attending treatment and district schools in the 2016-17 academic year. As shown in Table 65, BellXcel scholars reported a significantly more negative attitude toward learning when compared to students attending treatment schools ($p < 0.10$) and district schools ($p < 0.10$) in the 2016-17 academic year.

Table 65. Exploratory Outcome Unmatched Regression Results for Student Attitude toward Learning Fall 2016

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched – Regression (n = 452)	Effect Size	Unmatched – Regression (n = 912)	Effect Size
Valuing Education	**-0.13†** (0.07)	-0.27	**-0.13†** (0.07)	-0.27

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 390.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared student attitude toward learning between BellXcel scholars and students attending treatment and district schools in the 2017-18 academic year. As shown in Table 66, there were no significant differences between BellXcel scholars and students attending treatment and district schools in 2017-18 on outcomes in their self-reported attitude toward learning.

Table 66. Exploratory Outcome Unmatched Regression Results for Student Attitude toward Learning Fall 2017

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched – Regression (n = 445)	Effect Size	Unmatched – Regression (n = 1,108)	Effect Size
Valuing Education	-0.05 (0.07)	-0.10	-0.04 (0.06)	-0.08

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared the quality of students' self-reported relationships with adults in their school between BellXcel scholars and students attending treatment and district schools in the 2016-17 academic year. As shown in Table 67, there were no significant differences in students' self-reported relationships with teachers and adults in their school between BellXcel scholars and students attending treatment and district schools during the 2016-17 academic year.

Table 67. Exploratory Outcome Unmatched Regression Results for Relationships with Caring Adults Fall 2016

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched – Regression (n = 452)	Effect Size	Unmatched – Regression (n = 912)	Effect Size
Relationships with Teachers	-0.08 (0.09)	-0.13	0.07 (0.09)	0.10
Relationships with Caring Adults	-0.06 (0.09)	-0.10	0.03 (0.08)	0.04

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 390.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared the quality of students' self-reported relationships with adults in their school between BellXcel scholars and students attending treatment and district schools in the 2017-18 academic year. As shown in Table 68, there were no significant differences in students' self-reported relationships with teachers and adults in their school between BellXcel scholars and students attending treatment and district schools during the 2017-18 academic year.

Table 68. Exploratory Outcome Unmatched Regression Results for Relationships with Caring Adults Fall 2017

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched – Regression (n = 445)	Effect Size	Unmatched – Regression (n = 1,108)	Effect Size
Relationships with Teachers	-0.11 (0.08)	-0.17	0.02 (0.07)	0.03
Relationships with Caring Adults	-0.07 (0.08)	-0.11	0.02 (0.08)	0.03

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared outcomes on school engagement between BellXcel scholars and students attending treatment and district schools in the 2016-17 academic year. As shown in Table 69, there were no significant differences in students' self-reported school engagement or school belonging between BellXcel scholars and students attending treatment and district schools during the 2016-17 academic year.

Table 69. Exploratory Outcome Unmatched Regression Results for School Fall 2016

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched Regression (n = 452)	Standard Mean Difference	Unmatched Regression (n = 912)	Standard Mean Difference
School Engagement	**-0.17†** (0.10)	-0.25	-0.06 (0.10)	-0.09
School Belonging	-0.02 (0.09)	-0.03	0.05 (0.09)	0.07

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 390.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

.

Researchers compared outcomes on school engagement between BellXcel scholars and students attending treatment and district schools in the 2017-18 academic year. As shown in Table 70, there were no significant differences in students' self-reported school engagement or school belonging between BellXcel scholars and students attending treatment and district schools during the 2017-18 academic year.

Table 70. Exploratory Outcome Unmatched Regression Results for School Engagement Fall 2017

	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched Regression (n = 445)	Effect Size	Unmatched Regression (n = 1,108)	Effect Size
School Engagement	-0.07 (0.08)	-0.11	0.05 (0.07)	0.07
School Belonging	-0.13 (0.09)	-0.18	-0.05 (0.08)	-0.07

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared the frequency with which students participated in a variety of activities during the summer between BellXcel scholars and students attending treatment and district schools during the 2016-17 academic year. As shown in Table 71, BellXcel scholars did not participate in any of the measured activities at a significantly different rate than students attending treatment schools during the 2016-17 academic year. However, when compared to students attending district schools, BellXcel scholars reported going to the library ($p < 0.10$) and playing math games ($p < 0.001$) significantly more during the summer of 2016. Further, BellXcel scholars reported playing on their phones, watching TV, playing video games ($p < 0.10$) and going to a YMCA or community center($p < 0.05$) at a significantly lower frequency in the summer of 2016 than students attending district schools during the 2016-17 academic year.

Table 71. Exploratory Outcome Unmatched Regression Results for Frequency of Summer Activities Fall 2016

How often did you…	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched Regression (n = 452)	Effect Size	Unmatched Regression (n = 912)	Effect Size
…go to the library?	0.20 (0.26)	0.11	**0.47†** (0.26)	0.26
…write something?	0.10 (0.21)	0.05	0.29 (0.20)	0.16
…play math games?	0.41 (0.25)	0.22	**1.15*** ** (0.24)	0.63
…read a book?	-0.24 (0.21)	-0.13	-0.16 (0.18)	-0.09
…play on your phone, watch TV, or play video games?	0.13 (0.26)	-0.07	**-0.45†** (0.26)	-0.25
…go to community center, YMCA, or camp?	0.02 (0.22)	0.01	**-0.47*** (0.22)	-0.26
…play outside?	-0.06 (0.23)	-0.03	-0.22 (0.22)	-0.12

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. BellXcel vs. Treatment N = 426; BellXcel vs. District N = 390.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Researchers compared the frequency with which students participated in a variety of activities during the summer between BellXcel scholars and students attending treatment and district schools during the 2016-17 academic year. As shown in Table 72, when compared to students attending treatment schools, BellXcel scholars reported writing something ($p < 0.01$) and going to a community center or YMCA ($p < 0.01$) at a significantly higher frequency during the summer of 2017. When compared to students attending district schools, BellXcel scholars reported writing something ($p < 0.001$) and playing math games ($p < 0.05$) significantly more during the summer of 2017. However, BellXcel scholars reported playing outside ($p < 0.10$) at a significantly lower frequency in the summer of 2016 than students attending district schools during the 2016-17 academic year.

Table 72. Exploratory Outcome Unmatched Regression Results for Frequency of Summer Activities Fall 2017

How often did you…	BellXcel vs. Comparison Students: Treatment Schools		BellXcel vs. Comparison Students: District Schools	
	Unmatched Regression (n = 445)	Effect Size	Unmatched Regression (n = 1,105)	Effect Size
…go to the library?	0.16 (0.22)	0.09	0.32 (0.22)	0.18
…write something?	**0.61**** (0.24)	0.34	**0.89***** (0.26)	0.49
…play math games?	0.16 (0.24)	0.09	**0.57*** (0.22)	0.31
…read a book?	0.32 (0.23)	0.18	0.19 (0.20)	0.11
…play on your phone, watch TV, or play video games?	-0.21 (0.24)	-0.12	-0.38 (0.24)	-0.20
…go to community center, YMCA, or camp?	**0.61**** (0.23)	0.34	0.27 (0.21)	0.15
…play outside?	-0.25 (0.24)	-0.14	**-0.44†** (0.23)	-0.25

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. BellXcel vs. Treatment N = 444; BellXcel vs. District N = 468.
Note: Table presents the regression coefficients of the BellXcel variable and robust SE in parentheses from multivariate regressions that also control for student demographics, grade, and pretreatment outcome measure.

Appendix K. Additional Tables

Figure 18. STAR Pre-test and Post-test Scale Scores in Math Summer 2016

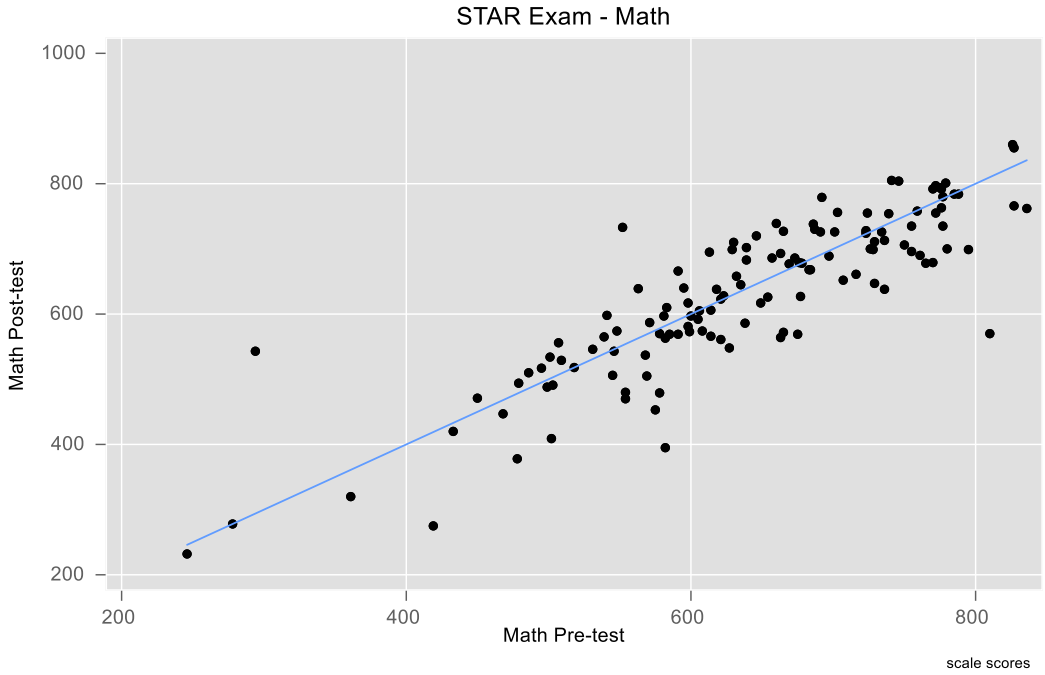
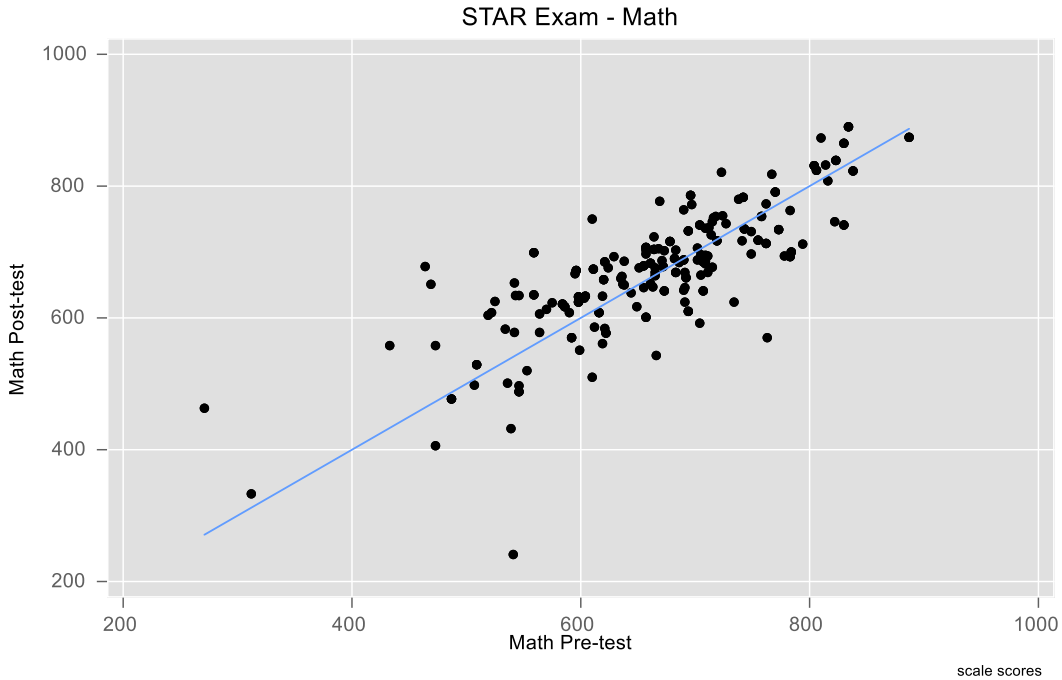


Figure 19. STAR Pre-test and Post-test Scale Scores in Math Summer 2017



References

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007a). Lasting Consequences of the Summer Learning Gap. *American Sociological Review*, *72*(2), 167–180. https://doi.org/10.1177/000312240707200202

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007b). Summer learning and its implications: insights from the Beginning School Study. *New Directions for Youth Development*, (114), 11–32. https://doi.org/10.1002/yd.210

American Institutes of Research. (2015). Are You Ready to Assess Social and Emotional Development? Tools Index. Retrieved from https://www.air.org/sites/default/files/AIR%20Ready%20to%20Assess_ACT_rev.pdf

Augustine, C. H., McCombs, J. S., Pane, J. F., Schwartz, H. L., Schweig, J., McEachin, A., & Siler-Evans, K. (2016). *Learning from Summer: Effects of Voluntary Summer Learning Programs on Low-Income Urban Youth*. Santa Monica, CA: RAND Corporation.

Augustine, Catherine H., McCombs, J. S., Schwartz, H. L., & Zakaras, L. (2013). Getting to Work on Summer Learning [Product Page]. Retrieved November 14, 2017, from https://www.rand.org/pubs/research_reports/RR366.html

Balfanz, R., & Byrnes, V. (2012). *The Importance of Being in School: A Report on Absenteeism in the Nation's Public Schools - SI&A - School Innovations & Achievement*. Retrieved from https://www.sia-us.com/news-and-events/sia-research-articles/the-importance-of-being-in-school-a-report-on-absenteeism-in-the-nations

Balfanz, R., Herzog, L., & Iver, D. J. M. (2007). Preventing Student Disengagement and Keeping Students on the Graduation Path in Urban Middle-Grades Schools: Early Identification and Effective Interventions. *Educational Psychologist*, *42*(4), 223–235.

Bifulco, R. (2012). Can Nonexperimental Estimates Replicate Estimates Based on Random Assignment in Evaluations of School Choice? A Within-Study Comparison. *Journal of Policy Analysis and Management*, *31*(3), 729–751. https://doi.org/10.1002/pam.20637

Birmingham, J., Pechman, E. M., Russell, C. A., & Mielke, M. (2005). *Shared Features of High-Performing After-School Programs: A Follow-Up to the TASC Evaluation*. Policy Studies Associates, Inc. Retrieved from https://eric.ed.gov/?id=ED498790

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. West Sussex, UK: John Wiley & Sons.

Borman, G. D., Benson, J., & Overman, L. T. (2005). Families, Schools, and Summer Learning. *The Elementary School Journal*, *106*(2), 131–150. https://doi.org/10.1086/499195

Borman, G. D., & Dowling, N. M. (2006). Longitudinal Achievement Effects of Multiyear Summer School: Evidence From the Teach Baltimore Randomized Field Trial. *Educational Evaluation and Policy Analysis*, *28*(1), 25–48. https://doi.org/10.3102/01623737028001025

Chaplin, D., & Capizzano, J. (2006). Impacts of a Summer Learning Program. Retrieved November 14, 2017, from https://www.urban.org/research/publication/impacts-summer-learning-program

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.

Cohen, M., Halfacre, A., Kearns, S., Knouse, S., Quinn, J., & Winiski, M. (2017). *Focused Needs and Asset Assessment*. Furman University: Furman University.

Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management*, *27*(4), 724–750. https://doi.org/10.1002/pam.20375

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review. *Review of Educational Research*, *66*(3), 227–268. https://doi.org/10.3102/00346543066003227

Corrin, W., Parise, L. M., Cerna, O., Haider, Z., & Somers, M.-A. (2015). *Case Management for Students at Risk of Dropping Out: Implementation and Interim Impact Findings from the Communities in Schools Evaluation*. MDRC. Retrieved from https://eric.ed.gov/?id=ED558497

Corrin, W., Sepanik, S., Rosen, R., & Shane, A. (2016, June 20). Addressing Early Warning Indicators [Text]. Retrieved June 13, 2017, from http://www.mdrc.org/publication/addressing-early-warning-indicators

Entwisle, D. R., & Alexander, K. L. (1992). Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School. *American Sociological Review*, *57*(1), 72–84. https://doi.org/10.2307/2096145

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fortson, K., Verbitsky-Savitz, N., Kopa, E., & Gleason, P. (2012). *Using an Experimental Evaluation of Charter Schools to Test Whether Nonexperimental Comparison Group Methods Can Replicate Experimental Impact Estimates. NCEE 2012-4019*. National Center for Education Evaluation and Regional Assistance. Retrieved from https://eric.ed.gov/?id=ED531481

Goldschmidt, P., & Wang, J. (1999). When Can Schools Affect Dropout Behavior? A Longitudinal Multilevel Analysis. *American Educational Research Journal*, *36*(4), 715–738. https://doi.org/10.3102/00028312036004715

Guo, S., & Fraser, M. W. (2014). *Propensity Score Analysis: Statistical Methods and Applications*. Thousand Oaks, CA: Sage Publications.

Harwell, M., & LeBeau, B. (2010). Student Eligibility for a Free Lunch as an SES Measure in Education ResearchEducational Researcher - Michael Harwell, Brandon LeBeau, 2010. Retrieved from http://journals.sagepub.com/doi/abs/10.3102/0013189X10362578

Imbens, G. W., & Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, *47*(1), 5–86. https://doi.org/10.1257/jel.47.1.5

Laird, J., Kienzi, G., DeBell, M., & Chapman, C. (2007). *Dropout rates in the United States: 2005:* Washington, D.C: U.S. Department of Education.

Lipsey, M. W., & Wilson, D. (2000). *Practical Meta-Analysis*. Thousand Oaks, Calif: SAGE Publications, Inc.

McCombs, J. S., Crego, A., Gershwin, D., Ikemoto, G. S., Le, V.-N., Kirby, S. N., … Xia, N. (2009). Ending Social Promotion Without Leaving Children Behind [Product Page]. Retrieved November 14, 2017, from https://www.rand.org/pubs/monographs/MG894.html

Montgomery, J. M., Nyhan, B., & Torres, M. (2018). How conditioning on posttreatment variables can ruin your experiment and what to do about it. *American Journal of Political Science*, *62*(3), 760–775.

Northwest Evaluation Association. (2004). *Reliability and Validity Estimates: NWEA Achievement Level Tests and Measures of Academic Progress*. Lake Owesgo, OR.

Perugini, M., Gallucci, M., & Costantini, G. (2018). A Practical Primer To Power Analysis for Simple Experimental Designs. *International Review of Social Psychology*, *31*(1), 20. https://doi.org/10.5334/irsp.181

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*. National Center for Education Evaluation and Regional Assistance. Retrieved from https://eric.ed.gov/?id=ED511781

Putnam, R. F., Luiselli, J. K., Handler, M. W., & Jefferson, G. L. (2003). Evaluating Student Discipline Practices in a Public School Through Behavioral Assessment of Office Referrals. *Behavior Modification*, *27*(4), 505–523. https://doi.org/10.1177/0145445503255569

Reardon, S. F., & Robinson, J. P. (2007). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In H. A. Ladd & E. B. Fiske (Eds.), *Handbook of Research in Education Finance and Policy*. Lawrence Erlbaum. Retrieved from https://cepa.stanford.edu/content/patterns-and-trends-racialethnic-and-socioeconomic-academic-achievement-gaps

Reschly, A. L., & Christenson, S. L. (2006). Prediction of Dropout Among Students With Mild Disabilities: A Case for the Inclusion of Student Engagement Variables. *Remedial and Special Education*, *27*(5), 276–292. https://doi.org/10.1177/07419325060270050301

Rockman et al. (2013). *Student Outlook Survey: Attitudes Toward Learning* (p. 15).

Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.2307/2335942

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, *79*(387), 516–524. https://doi.org/10.2307/2288398

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*, *39*(1), 33–38. https://doi.org/10.2307/2683903

Rumberger, R. W., & Lim, S. A. (2008). *Why Students Drop Out of School: A Review of 25 Years of Research* (California Dropout Research Project Report #15). Retrieved from https://www.issuelab.org/resource/why-students-drop-out-of-school-a-review-of-25-years-of-research.html

Sanders, W. L., & Horn, S. P. (1998). Research Findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for Educational Evaluation and Research. *Journal of Personnel Evaluation in Education*, *12*(3), 247–256. https://doi.org/10.1023/A:1008067210518

Schochet, P. Z. (2008). *Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations. NCEE 2008-4018*. National Center for Education Evaluation and Regional Assistance. Retrieved from https://eric.ed.gov/?id=ED501605

Somers, M.-A., Welbeck, R., Grossman, J., & Gooden, S. (2015, March 11). An Analysis of the Effects of an Academic Summer Program for Middle School Students [Text]. Retrieved November 14, 2017, from https://www.mdrc.org/publication/analysis-effects-academic-summer-program-middle-school-students-0

StataCorp. (2017). *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.

Steiner, P. M., & Cook, D. (2013). Matching and Propensity Scores. In T. D. Little (Ed.), *The Oxford Handbook of Quantitative Methods in Psychology* (Vol. 1, pp. 237–259). New York: Oxford University Press.

Stetser, M. C., & Stillwell, R. (2014). *Public High School Four-Year On-Time Graduation Rates and Event Dropout Rates: School Years 2010-11 and 2011-12. First Look. NCES 2014-391*. National Center for Education Statistics. Retrieved from https://eric.ed.gov/?id=ED544798

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *25*(1), 1–21. https://doi.org/10.1214/09-STS313

Stuart, E. A., & Rubin, D. B. (2008). Best Practices in Quasi–Experimental Designs: Matching Methods for Causal Inference. In J. Osborne, J. Osborne (Ed.), *Best Practices in Quantitative Methods* (pp. 155–176). 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc. https://doi.org/10.4135/9781412995627.d14

Sugai, G., Sprague, J. R., Horner, R. H., & Walker, H. M. (2000). Preventing School Violence: The Use of Office Discipline Referrals To Assess and Monitor School-Wide Discipline Interventions. *Journal of Emotional and Behavioral Disorders*, *8*(2), 94–101.

United States Department of Agriculture. (2015). Child Nutrition Programs: Income Eligibility Guidelines. *Federal Register*, *80*(61), 17026 – 17027.

Walker, D. (2018, April 29). As outsiders discover the city, Greenville has joined many top 10 and best-of lists. *Greenville News*.

Wang, S., McCall, M., Jiao, H., & Harris, G. (2012). *Construct Validity and Measurement Invariance of Computerized Adaptive Testing: Application to Measures of Academic Progress (MAP) Using Confirmatory Factor Analysis*. Retrieved from https://eric.ed.gov/?id=ED531495

What Works Clearinghouse. (2017). *Standards handbook (Version 4.0)*. Washington, D.C.: Institute of Education Sciences.