AIR®

AMERICAN INSTITUTES FOR RESEARCH®

New Jersey 21st Century Community Learning Centers Year 4 Evaluation Report Impact Data for 2015–16

Matthew Vinson
Andrew Swanlund

NOVEMBER 2017

Funding Statement

This project was funded in its entirety from the federal Elementary and Secondary Education Act (ESEA), Title IV, Part B, 21st Century Community Learning Centers (21st CCLC) grant through a contract with the New Jersey Department of Education. The ESEA was reauthorized in 2015 by the Every Student Succeeds Act (ESSA).

New Jersey 21st Century Community Learning Centers Year 4 Evaluation Report Impact Data for 2015–16

November 2017

Matthew Vinson
Andrew Swanlund

AMERICAN INSTITUTES FOR RESEARCH®

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

www.air.org

Copyright © 2017 American Institutes for Research. All rights reserved.

3067_11/17

Contents

Page

Executive Summary i
- 2015–16 Context i
- Data Sources ii
- Analysis ii
- Summary of Key Findings iii
- Staff Survey Data iv
- Program Impact Estimates v
- Conclusions and Next Steps vi

Chapter 1. Introduction 1

Chapter 2. Evaluation Approach, Limitations, and Conceptual Framework 3
- Approach 3
- Limitations and Challenges 5
- AIR's Conceptual Framework 6

Chapter 3. Grantee, Center, and Student Characteristics 13
- Grantee Characteristics 13
- Key Center Characteristics 15
- Student Characteristics 20

Chapter 4. Staff Survey Data 26
- Staff Survey Response Rates and Respondent Population Note 27
- Survey Responses by Survey Area 28
- Summary of Staff Survey Data 37

Chapter 5. Assessing 21st CCLC Program Outcomes 39
- Results 41
- Summary of Results 48

Chapter 6. Conclusions and Next Steps 50

References 51

Appendix A. New Jersey 21st CCLC Staff Survey A-1

Appendix B. Variables Used in Propensity Score Matching Models B-1

Tables

Page

Table 1. Overall Statistics on Number of Center Staff 16

Table 2. Average Student–Teacher Ratio per Center, 2015–16 17

Table 3. Summary of Demographic Information for Students, 2015–16 20

Table 4. Continuous Years of Student Participation, 2015–16 22

Table 5. Total School-Year Hours of Attendee Participation, by Activity Type 23

Table 6. Average Number of Hours in Reading and Mathematics per Student, 2015–16 25

Table 7. Responses Concerning Program Goals and Vision 28

Table 8. Responses Concerning Collective Staff Efficacy 29

Table 9. Responses Concerning Provision of Opportunities for Youth Autonomy and Choice 30

Table 10. Responses Concerning Program Design 31

Table 11. Responses Concerning Communication and Linkages to the School Day 32

Table 12. Responses Concerning Data Use 34

Table 13. Responses Concerning Service Delivery Practices 34

Table 14. Responses Concerning Youth Ownership 35

Table 15. Responses Concerning Internal Communication 36

Table 16. Responses Concerning Parent Communication 37

Table 17. Pooled Impact of Participation on Truancy and Grade Retention 42

Table 18. Grade-Level Impact of Participation on Truancy 43

Table 19. Grade-Level Impact of Participation on Grade Retention 43

Table 20. Pooled Impact of Participation on Mathematics and Reading Achievement 44

Table 21. Grade-Level Impact of Participation on Mathematics Achievement 45

Table 22. Grade-Level Impact of Participation on Reading Achievement 45

Table 23. Pooled Impact of High Attendance on Truancy and Grade Retention (60+ days versus 1 to 29 days) 46

Table 24. Grade-Level Impact of High Attendance on Truancy (60+ days versus 1 to 29 days) 46

Table 25. Grade-Level Impact of High Attendance on Grade Retention (60+ days versus 1 to 29 days) 47

Table 26. Pooled Impact of High Attendance on Mathematics and Reading Achievement (60+ days attendance and 60+ hours subject-specific programming versus 1 to 29 days attendance and 30 or fewer hours subject-specific programming) 47

Table 27. Grade-Level Impact of High Attendance on Mathematics Achievement (60+ days attendance and 60+ hours mathematics programming versus 1 to 29 days attendance and 30 or fewer hours mathematics programming) 47

Table 28. Grade-Level Impact of High Attendance on Reading Achievement (60+ days attendance and 60+ hours reading programming versus 1 to 29 days attendance and 30 or fewer hours reading programming) 48

Table 29. Relationship Between 21st CCLC Program Attendance and Student Outcomes for Participating Students With an IEP 48

Figures

Page

Figure 1. How Change Happens in 21st CCLC Programming 7

Figure 2. Evaluation and the Theory of Change 11

Figure 3. Number of Grantees by Year of Operation, 2014–15 and 2015–16 14

Figure 4. Number of Grantees by Organization Type 14

Figure 5. Total Number of Staff by Staff Type 16

Figure 6. Number of Centers by Grade Level Served 18

Figure 7. Percentage of Centers Offering Activities Linked to a Given Theme 19

Figure 8. Percentage of Total Activity Minutes Dedicated to Activity Themes, Among Centers With Each Theme 19

Figure 9. Number of Students Served in 21st CCLC by Attendance Gradation 21

Figure 10. Proportion of Time Each Participant Spends on Activities of a Given Type (Averages) 23

Figure 11. Percentage of Attendees With 10 or More Hours in a Given Activity Type (School Year) 24

Figure 12. Level of Education Completed by 21st CCLC Staff Responding to the Staff Survey 27

Figure 13. Staff Roles in 21st CCLC Programming Among Survey Respondents 28

Executive Summary

The information presented in this report is the result of data collected and analyzed as part of a statewide evaluation of New Jersey's 21st Century Community Learning Centers (21st CCLC) program, currently being conducted by American Institutes for Research (AIR). The results outlined in this report are associated with 21st CCLC–funded activities and services delivered during the course of the 2015–16 school year, though some findings from 2014–15 offer cross-year comparisons. Such comparisons are present only when they are noteworthy or particularly useful.

The information collected and analyzed in relation to the 2015–16 school year was meant to answer three primary evaluation questions related to the implementation of the New Jersey 21st CCLC program:

1. What were the primary characteristics of programs funded by 21st CCLC and the students served?
2. How many youth with individual education plans (IEPs) were served by the program, and what outcome levels are associated with their participation in the 21st CCLC program in terms of mathematics and reading assessments, truancy, and retention?
3. To what extent is there evidence that students participating in 21st CCLC program services and activities demonstrate better outcomes compared with students not participating in the program, specifically with respect to:
    a. Higher academic achievement in reading/language arts and mathematics
    b. Lower truancy and retention rates

Note that there is normally another question concerning state-level performance on program leading indicators; because the leading indicators were not calculated for 2015–16 (in agreement with New Jersey Department of Education [NJDOE]), these values are not in this report. As was stated in the previous year's report, indicator values will be reported next year, following the new, revised format.

2015–16 Context

The broader evaluation context is one of finalizing transition. Over the course of 2015–16 and 2016–17, a great deal of previous data collection was reassessed and revised, with AIR working closely with NJDOE and the Evaluation Advisory Group to refine and streamline data components to better support quality improvement, monitoring, and impact assessment. These efforts included:

- Revision of 21st CCLC program leading indicators, cutting back from 22 to 12
- Removal of the Evaluation Tracking and Reporting System (ETRS) end-year data collection (entirely)
- Revision of the ETRS midyear data collection, dividing collection into two time points and reducing question burden
- Introduction of a new self-assessment module within ETRS

- Introduction of a new action-research module within ETRS
- Introduction of the youth motivation, engagement, and beliefs survey

To implement these changes, standard data collection using ETRS was paused during 2015–16 (i.e., data collection for the end-year and midyear data collections). This provided the AIR evaluation team with time to make system changes (which were realized prior to the 2016–17 data collection year). Consequently, certain types of data (e.g., leading indicator data) are not presented in this report, as indicated. Also, youth survey data are not included in this report; though youth survey data collection is currently underway, a pilot of the new youth surveys was deployed only in spring 2016, and those pilot results were already reported in the year three report. The impact report produced next year, for 2016–17 data, will include outcome data from the youth surveys as part of the impact models.

Data Sources

To address the aforementioned evaluation questions, data were collected from the following sources:

- **Program Activity and Review System (PARS21).** PARS21 is a web-based data collection system developed and maintained by NJDOE that collects directly from grantees a broad array of program-characteristic, student-demographic, attendance, and outcome data throughout the program year.
- **Staff Survey.** The purpose of the online staff survey was to obtain information from staff members working directly with youth in programs funded by 21st CCLC about the extent to which they engage in practices suggested by the afterschool research literature as likely to be supportive of both positive academic and youth development outcomes. The staff survey is collected via AIR's Evaluation Tracking and Reporting System (ETRS).
- **New Jersey Standards Measurement and Resource for Teaching (NJ SMART) Data Warehouse.** In early 2017, the research team obtained access to New Jersey assessment test scores, truancy data, and grade promotion data for the 21st CCLC participants served during the 2015–16 school year. These data came from the NJ SMART data warehouse maintained by NJDOE for students in Grades 4–8. Similar data also were obtained for students attending the same schools as the 21st CCLC participant population who did not participate in the program during these periods. The research team used these data to conduct an analysis of the program impact outcomes.

Analysis

The findings in this report are purely quantitative, with methods as follows:

1. *Descriptive Analyses.* Data related to grantee, center, and student characteristics obtained from PARS21, NJ SMART, and the staff survey were analyzed descriptively to explore the range of variation on a given characteristic.
2. *Correlational Multilevel Modeling Techniques*. As part of a set of analyses based on NJDOE interest, a multilevel model was run to explore the relationship between participation levels (in terms of days) and student outcomes in which students were

identified as having an IEP. Outcomes analyzed included performance on state assessments in reading and mathematics, truancy, and grade promotion. Although these analyses afford the capacity to identify correlations between participation level and the outcomes assessed, the method is not sufficient to indicate cause.

3. *Propensity Score Matching*. In contrast to the multilevel modeling techniques just described, propensity score matching approaches were employed to estimate the causal impact of 21st CCLC program participation on student performance in terms of truancy rates, grade promotion, and achievement (reading and mathematics). Given that 21st CCLC program participants were not randomly assigned to participate in the program, the problem of selection bias was an issue that needed to be addressed before program impact could be explored from a causal perspective. It is likely that students who participated in 21st CCLC programming were different from those students attending the same schools who did not enroll in 21st CCLC. These differences can bias estimates of program effectiveness because they make it difficult to disentangle preexisting differences between participants and nonparticipants from program impact. Propensity score matching was used to mitigate that existing selection bias in program effect.

Multiple analyses were conducted using propensity score matching in particular.

Summary of Key Findings

The following is a summary of key evaluation findings.

Primary Characteristics of Programs Funded by 21st CCLC and the Students Served

Grantee Characteristics

- A plurality of grantees (44%) were in their second year of program operation.
- Grantees were split between the categories of school-based (44 percent) and non-school-based (56 percent) grantees (about the same as in 2014-15).

Center Characteristics

- By far, the most common staff type reported by grantees was school-day teacher, with 845 (39% of all staff) being reported for the 2015–16 school year. The next highest category was "program staff"[1] with 546 (or 25% of all staff).
- Centers on average had 18 staff members (median 16).
- The average student-to-staff ratio was about 11 students for each program staff member, similar to the ratio observed in 2014–15.
- Centers mainly served children in elementary and middle schools exclusively (82% of centers, the same as the previous year).

[1] "Program staff" is a category of staff reported in PARS21.

- Approximately 28% of all centers chose career awareness as their theme, whereas another 43% chose science, technology, engineering, and mathematics (STEM). Another 20% chose visual and performing arts as their central theme, and only 4% chose civic engagement.

Student Characteristics

- A total of 15,449 students attended 21st CCLC programming for at least one day (compared with 15,049 for 2014–15).
- A majority of 21st CCLC participants were Hispanic/Latino (45%) or Black (35%). Most attendees (75%) qualified for free or reduced-price lunch.
- Two thirds of the students (67.2%) attended 30 or more days, and slightly more than one third (35.3%) participated for 90 or more days.
- About 31% of students attended 21st CCLC programming for two consecutive years or more.
- On average, students spent about 20% of their time in tutoring or homework help, about 19% in academic enrichment, 18% in youth development/learning activities, and 16% in recreation.

  However, taking the *median* total student hours spent in each type of activity (instead of the average) showed that students spent a median of 12 hours in academic enrichment. This was followed by nine hours in youth development/learning activities, and eight hours in recreational activities. The median total number of hours was two. This indicates that the averages (as indicated in the preceding bullet point) were skewed by a smaller number of youth with high amounts of tutoring.
- A total of 52% of all youth participated in at least 10 hours of academic enrichment across the year. Comparable figures for youth development/learning activities, recreation, and tutoring were 49%, 48%, and 44%, respectively.
- The typical student attended an average of 54 hours of reading activities and 50 hours of mathematics activities (average of total hours across the reporting period). This was significantly higher than the previous year but was likely driven by a change in the method of activity reporting whereby grantees can report activities as targeting multiple subjects (instead of one subject).

Staff Survey Data

Overall, 546 program staff survey responses were received in December 2015 and January 2016. Of these, 530 were completed entirely. These surveys were received from 67 centers associated with 30 grantees, with an average of 8.1 and a median of 8.0 surveys per center. Many of the newest grantees were not included in the staff survey (per NJDOE guidance), nor were grantees that had just concluded programming, so this number, despite being significantly lower than the overall grantee count, does not indicate a low response rate (which was actually 92%). What this does mean, however, is that the survey results are not necessarily indicative of all grantee staff opinions but are instead the responses of staff working at programs not considered “new” by NJDOE.

Overall, respondents tended to be well-educated (with nearly all having at least some college education and approximately four of every five staff having at least a college degree) and tended to respond positively on most if not all scales. Respondents also strongly indicated that provision of enrichment opportunities was a primary goal (with nearly 90% selecting this as a primary goal of their program), and staff generally reported high levels of staff efficacy (e.g., staff actively consulting youth, providing structured and planned activities).

Nevertheless, there were areas that staff indicated could be improved. Respondents indicated less enthusiastic *agreement* about whether staff provide youth with ongoing opportunities for reflection (51% indicating simply "agree" where "strongly agree" was the highest option) and were somewhat more likely to respond "rarely" (2.8%) or "sometimes" (22.0%) concerning participating in or leading activities structured to respond to youth feedback. In this same vein, approximately a quarter of respondents indicated that youth do not have opportunities to determine how content is covered in program offerings (25.7%, "strongly disagree" and "disagree" combined) or to help create rules and guidelines for the program (23.2%, "strongly disagree" and "disagree" combined). Of course, staff did report tying activities to specific learning goals, and staff generally indicated that activities that enable collaborative work or types of youth leadership are provided at least occasionally.

In addition, roughly a quarter of respondents indicated they do *not* meet regularly with school-day staff to discuss academic progress, and questions concerning communication with the school day yielded many answers in the "agree" column (rather than "strongly agree"). Frequent ("often") use of data such as academic plans, test scores, grades, and teacher input was reported by only a minority of respondents. Furthermore, whereas internal program communication was reported as generally strong (with approximately three quarters of respondents indicating modestly frequent communication with peers or team members, depending on communication type), a large minority of staff members (38%) indicated that they never observe other staff delivering activities (for feedback purposes). In addition, a full quarter of staff indicated their program "never" asks for parent feedback. Noting again that these responses were provided by existing grants, this finding indicates some possible areas for professional development.

Program Impact Estimates

Program impact analysis was conducted using propensity score matching to create comparison groups, analyzing three different types of outcomes: truancy rates, retention in grade level, and assessment scores. Two different types of comparison were performed: participant versus nonparticipant (but attending the same schools) and high-attending 21st CCLC participant (i.e., 60 or more days) versus low-attending 21st CCLC participant (one to 29 days). The propensity score stratification approach seeks to minimize the impact of selection bias on the estimates of program impact; the second type of analysis, low-attending 21st CCLC participants versus high-attending 21st CCLC participants, was intended to further reduce selection bias. Of course, it is an untestable assumption that such models can fully account for selection bias.

Concerning participant versus nonparticipant comparisons, the impact analysis results indicate that 21st CCLC programming has a positive impact on truancy rates for the 60-or-more-days group, though the effect size is very small. The impact of programming on truancy rates is larger for Grades 6–8, with rate ratios of .848, .896, and .930, respectively, for youth attending above

the 60-day threshold. The effect of programming on truancy rates for Grades 6 and 8 was also statistically significant at the 30-day threshold, though small.

The overall impact of 21st CCLC programming on retention was apparent at the 60-or-more-days level, with an effect size odds ratio of .876. Looking at retention by grade levels, significant reductions were seen for Grade 6 students attending 30 or more days or 60 or more days (odds ratios of .441 and .270, respectively), whereas for Grade 7 there was a significant difference based on attending 60 or more days (odds ratio of .876). In Grade 8, the difference was for students who attended at least 30 days (odds ratio .240). Note, however, that the absolute incidence rate of retention was, on the whole, relatively low.

In terms of assessment score outcomes, there was evidence of a program effect on mathematics at the overall level, with about a .02 standard deviation unit difference between treatment and nontreatment groups at both the 30- and 60-or-more-days levels. This overall effect seemed to have been driven by effects witnessed for Grades 4 and 5 in particular, where the effect sizes were comparatively large (effect sizes between .07 and .09 standard deviation units). For context, note that Hill, Bloom, Black, and Lipsey (2008) found that, on average, the effect of a whole year of learning on assessment results (counting time both in and out of school) averaged 0.31 standard deviation units for reading and 0.42 standard deviation units for mathematics. There was, however, also a statistically significant *negative* effect observed for Grade 8 attendees at the 30-or-more-days level (though the overall *n* count for the treatment group was, in this case, relatively small, at a few hundred scores). No effects on reading scores were detected, in keeping with results of past impact analyses conducted for NJDOE.

The analyses on high attenders versus low attenders did not yield findings of particular interest, though there were statistically significant findings regarding mathematics for Grades 4 and 8 (positive in the case of the former, negative in the case of the latter). It may be the case that the attendance levels of the two groups (29 days and 60 or more days) may not be sufficiently spread apart for statistical models to detect patterns of difference—at least not with confidence sufficient to rule out chance. This is presently an untestable hypothesis, however, because pushing the attendance brackets further apart yields ever-smaller *n* sizes, which erodes analytic power.

Overall, the results are roughly in keeping with what has been observed in previous impact analyses conducted for NJDOE.

Conclusions and Next Steps

Overall, the 21st CCLC program in New Jersey seems to be serving the intended population and is offering activities in keeping with New Jersey's 21st CCLC goals. In terms of staff surveys, respondent answers indicate that there may be opportunities for NJDOE to offer professional development guidance. Providing more opportunities for youth to determine content, enabling staff to observe each other (with the goal of helping each other learn), and increasing feedback requests from parents are all areas in which staff could potentially use support. Asking grantees about interest in these areas might be a useful next step.

The program continues to have a modest if positive impact on truancy and retention, along with an observable impact on mathematics scores (though not for reading). The results indicate a need

for more sensitive outcome measures, however, which means the youth survey data collected for 2016–17 will be of high interest. Impact analyses conducted next year will incorporate the youth outcome measures.

Chapter 1. Introduction

For over 15 years, 21st CCLC, operating across New Jersey, has provided youths in high-poverty communities the opportunity to participate in academic enrichment programs and other development and support activities designed to enhance these youths' academic well-being. The primary purpose of this report, one in a series of evaluation reports, is to examine the impact of the New Jersey 21st CCLC programs, especially in terms of impact on youth outcomes of interest. This report provides a descriptive picture of the programs as well as a view of the impact results in context.

The information presented in this report is the result of data collected and analyzed as part of a statewide evaluation of New Jersey's 21st CCLC program, which AIR is currently conducting. For the most part, the results outlined in this report are associated with 21st CCLC–funded activities and services delivered during the 2015–16 school year, though some findings from 2014–15 provide cross-year comparisons. These comparisons are made only when cross-year comparison is useful.

The broad evaluation context is to finalize transition. Throughout 2015–16 and 2016–17, a great deal of previous data collection was reassessed and revised, with AIR working closely with NJDOE and the Evaluation Advisory Group to refine and streamline data components to better support quality improvement, monitoring, and impact assessment. These efforts included the following:

- Revision of 21st CCLC program leading indicators, cutting back from 22 to 12
- Removal of the Evaluation Tracking and Reporting System (ETRS) end-of-year data collection (entirely)
- Revision of the ETRS midyear data collection, dividing collection into two time points and reducing the question burden
- Introduction of a new self-assessment module within ETRS
- Introduction of a new action-research module within ETRS
- Introduction of the youth motivation, engagement, and beliefs survey

To implement these changes, standard data collection using ETRS was paused during 2015–16 (i.e., data collection for the end-year and midyear data collections). This provided the AIR evaluation team with time to make system changes (which were realized prior to the 2016–17 data collection year). Consequently, certain types of data are not presented in this report (such as leading indicator data). Also, youth survey data are not included in this report; a pilot of the new youth surveys was deployed only in spring 2016, and the pilot results were already reported in the previous year's report (for 2014–15). The impact report produced next year, for 2016–17 data, will include outcome data from the youth surveys as part of the impact models.

The information collected and analyzed in relation to the 2015–16 school year came from three primary sources: the staff surveys (collected during winter 2015–16), Program Activity and Review System (PARS21) data (collected directly from the grantees by NJDOE and

subsequently transferred to AIR), and NJDOE state warehouse data (which included student demographic data and state assessment scores). In this report, these data help address, as far as possible, three of the primary evaluation questions related to implementation of 21st CCLC programming in New Jersey:

1. What were the primary characteristics of programs funded by 21st CCLC and the students served?
2. How many youth with individual education plans (IEPs) received program services, and what outcome levels are associated with their participation in the 21st CCLC program in terms of mathematics and reading assessments, truancy, and retention?
3. To what extent is there evidence that students participating in 21st CCLC program services and activities have demonstrated better outcomes compared with students not participating in the program, specifically with respect to
    a. Higher academic achievement in reading/language arts and mathematics
    b. Lower truancy and retention rates

There is normally another question concerning state-level performance on program leading indicators; because the leading indicators were not calculated for 2015–16 (to free resources for ETRS system coding changes), these values are not presented in this report. As the previous year's report states, indicator values will be reported next year, following the new, revised format.

This report is organized as follows: Chapter 2 contains a summary of the evaluation questions and an explanation of why these questions are important to the field. It also includes a description of the analytic methods used to support the evaluation, including a description of data sources. Chapter 2 concludes with an overview of AIR's conceptual framework for how change happens in 21st CCLC programming. Chapter 3 provides an overview of grantee, site,[2] and youth characteristics, with a particular emphasis on characteristics that have been shown to be related to improving youth academic achievement and attaining desired program outcomes. Where appropriate, there are year-to-year comparisons (using 2014–15 data alongside 2015–16 data). Chapter 4 presents basic descriptive data for the 2015–16 staff survey. Chapter 5 provides impact results, focusing on truancy, grade promotion, and test score improvement. Chapter 6 concludes with next steps.

[2] In this report, the terms *site* and *program* are used to refer to the physical location where 21st CCLC–funded services and activities take place. Sites are characterized by defined hours of operation, have dedicated staffs, and usually have positions similar to site coordinators. Each 21st CCLC grantee in New Jersey has at least one site; many grantees have more than one site.

Chapter 2. Evaluation Approach, Limitations, and Conceptual Framework

This chapter presents an overview of AIR's evaluation approach, significant limitations, and the conceptual framework. The evaluation approach generally details the questions AIR attempted to answer through the course of the evaluation, along with data used to answer the questions and analytical techniques. The limitations section provides several caveats relating to the data sources themselves and to the findings in general. Finally, the chapter concludes with AIR's theory for how change happens in 21st CCLC programming, which provides a transition into the descriptive and impact data shown in the subsequent chapters.

Approach

For this year's evaluation report, AIR focused on three of the primary evaluation questions related to implementation of the New Jersey 21st CCLC program:

1. What were the primary characteristics of programs funded by 21st CCLC and the students served?
2. How many youth with IEPs were served by the program, and what outcome levels are associated with their participation in the 21st CCLC program in terms of mathematics and reading assessments, truancy, and retention?
3. To what extent is there evidence that students participating in 21st CCLC program services and activities demonstrate better outcomes compared with students not participating in the program, specifically with respect to the following:
    a. Higher academic achievement in reading/language arts and mathematics
    b. Lower truancy and retention rates

As Chapter 1 states, consideration of leading indicator performance (a question AIR has explored in the past) was not included this year because leading indicator values were not calculated for 2015–16. Given the nature of the questions, this report is fairly simple and straightforward in scope, keeping primarily to a basic presentation of program characteristics, a descriptive presentation of staff survey data, and, finally, an impact analysis of the outcomes of interest (as available).

As Chapter 1 mentions, the data used to explore the questions came from three different sources, which include information from administrative data systems and surveys. The following section outlines in greater detail each source and how it contributed to the project.

PARS21

PARS21 is a web-based data collection system that NJDOE developed and maintains to collect a broad array of program-characteristic, student demographic, attendance, and outcome data throughout the program year directly from grantees. Data extracted from PARS21 were used to construct variables summarizing the activity and staffing models employed by sites, program maturity and organization type, and levels of program attendance in relation to the 2015–16

school year. Data extracted from PARS21 used to carry out analyses summarized in this report are from fall 2016 and spring 2017.

Staff Survey

The purpose of the online staff survey was to obtain information from staff members working directly with youth in programs funded by 21st CCLC about the extent to which they engage in practices suggested by the afterschool research literature as likely to be supportive of both positive academic and youth development outcomes. Scales appearing on the survey included the following:

- Collective staff efficacy in creating interactive and engaging settings for youth;
- Intentionality in activity and session design;
- Practices supportive of academic skill building, including linkages to the school day and using data about student academic achievement to inform programming;
- Practices supportive of positive youth development;
- Opportunities for youth ownership;
- Staff collaboration and communication to support continuous program improvement; and,
- Practices supportive of parent involvement and engagement.

Staff members were selected as part of the survey sample if they were actively providing services at the site that directly served students participating in the program. The 21st CCLC project directors were instructed to select staff members who worked in their program the most frequently and delivered activities that were most aligned with their centers' objectives for student growth and development. The goal was to have project directors identify a minimum of 12 staff members per center to take the survey. In cases in which centers had fewer than 12 active staff members, all staff members working with students at the center were directed to take the survey. This data collection took place during December 2015 and January 2016. In all, complete surveys were obtained from 67[3] centers active during the 2015–16 school year, an average of approximately eight completed surveys per site. Appendix A contains the staff survey questions.

New Jersey Standards Measurement and Resource for Teaching Data Warehouse

In early 2017, the research team gained access to New Jersey assessment test scores, truancy, and grade promotion data for 21st CCLC participants served during the course of the 2015–16 school year. These data came from the New Jersey Standards Measurement and Resource for Teaching **(**NJ SMART) data warehouse maintained by NJDOE for students in Grades 4–8. Similar data also were obtained for students attending the same schools as the 21st CCLC participant population who did not participate in the program during these periods. These data were used to analyze the impact of the program outcomes.

[3] Many programs concluded during 2014–15; consequently, many programs had just started in 2015–16. Programs that were brand new were not required to administer the staff survey.

Analytic Approach and Methods

Although previous reports the AIR evaluation team prepared as part of this project have included findings predicated on both qualitative and quantitative approaches, the findings outlined in this report are purely quantitative. This approach was largely driven by the evaluation questions and by the data available to the evaluation team for 2015–16. Analyses highlighted in this report fall within three general categories:

1. *Descriptive Analyses.* Data related to grantee, center, and student characteristics obtained from PARS21, NJ SMART, and the staff survey were analyzed descriptively to explore the range of variation on a given characteristic.
2. *Correlational Multilevel Modeling Techniques*. As a set of analyses based on NJDOE interest, a multilevel model was run to explore the relationship between participation level (in terms of days) and student outcomes in which students were identified as having an IEP. Outcomes analyzed included performance on state assessments in reading and mathematics, truancy, and grade promotion. Although these analyses afford the capacity to identify correlations between participation level and the outcomes assessed, the method was not sufficient to indicate cause.
3. *Propensity Score Matching*. In contrast to the multilevel modeling techniques just described, propensity score matching approaches were employed to estimate the causal impact of 21st CCLC program participation on student performance in terms of truancy rates, grade promotion, and achievement (reading and mathematics). Given that 21st CCLC program participants were not randomly assigned to participate in the program, the problem of selection bias was an issue that needed to be addressed before program impact could be explored from a causal perspective. It is likely that students who participated in 21st CCLC programming were different from students attending the same schools who did not enroll in 21st CCLC. These differences could bias estimates of program effectiveness because they make it difficult to disentangle preexisting differences between participants and nonparticipants from program impact. Propensity score matching was used to mitigate that existing selection bias in program effect.

Multiple analyses were conducted using propensity score matching. The presentation of results includes more detail on each analytic approach.

Limitations and Challenges

It is important to note that there are limitations associated with the methods employed to support the evaluation. The primary limitation of the results highlighted in this report relate to the fact that most of the data sources are predicated (at least to some extent) on self-reported data provided by 21st CCLC grantee staff members. This characteristic of most of the data analyzed likely led to the introduction of some level of error into the process predicated on the following:

- **Imperfect Recall and Motivation.** The staff survey contains items that required respondents to mentally review events, conversations, practices, and experiences that took place during the course of a school year and then decide how best to report required data. It is likely that some respondents were more adept at remembering and analyzing than others were and that some responses were better than others were. Similarly,

PARS21 requires grantees to self-report attendance and activities, which may be done with more or less diligence (potentially related to each site's record-keeping habits). In addition, programs have been tasked by NJDOE to report state student identifiers in PARS21 (used to match each participant's data to NJ SMART records), which may have in some cases been reported incorrectly or incompletely. These are simply realities associated with self-reported data.

- **Social Desirability.** A limitation intrinsic to survey data is that, depending on question framing and individual psychology, respondents may have selected a response not because it was true but because they believed it was the answer they *should* give. For example, anyone reading the items appearing on each of the measures employed as part of a survey could quite easily select a response indicating a high level of program implementation function even though the response was not entirely true. Respondents motivated to put their program's best foot forward may even have inadvertently chosen a favorable response—one that reported a higher level of functioning than was the case—and thereby biased the estimate of 21st CCLC program implementation derived from their responses. This is simply a shortcoming of survey data.

Despite these potential error sources, the analyses in this report used the data available, and researchers attempted to continuously improve the quality of the data received. During the past several years, AIR and NJDOE staff have worked with grantees to emphasize the importance of submitting and maintaining high-quality data to help increase the overall accuracy and completeness of the evaluation data. This effort is ongoing.

As a final note on limitation, there is also one important caveat concerning the impact models: The models cannot control for variables unavailable at the time of the analysis. This is important because, in lieu of random assignment, propensity score matching uses available student and school demographics to build a comparison group. Therefore, nonparticipants are chosen to compose a comparison group such that the comparison group is *similar* to the treatment group (i.e., the 21st CCLC participants) in terms of demographic characteristics and prior-year performance levels. However, the models obviously cannot incorporate variables that do not exist or that were not provided to AIR. This means that if there are demographic characteristics that (a) are critical to determining selection for 21st CCLC programs and (b) are not available for use in comparison group construction, the resultant comparison group, with respect to that single variable, will be flawed: The comparison group may not end up being similar to the treatment group in that critical respect. This is why random assignment is considered a gold standard in research of this kind. However, when random assignment is not possible, creating a comparison group using the propensity score matching approach is the next best option available. Appendix B contains a list of the student- and school-level variables used in the propensity score matching process.

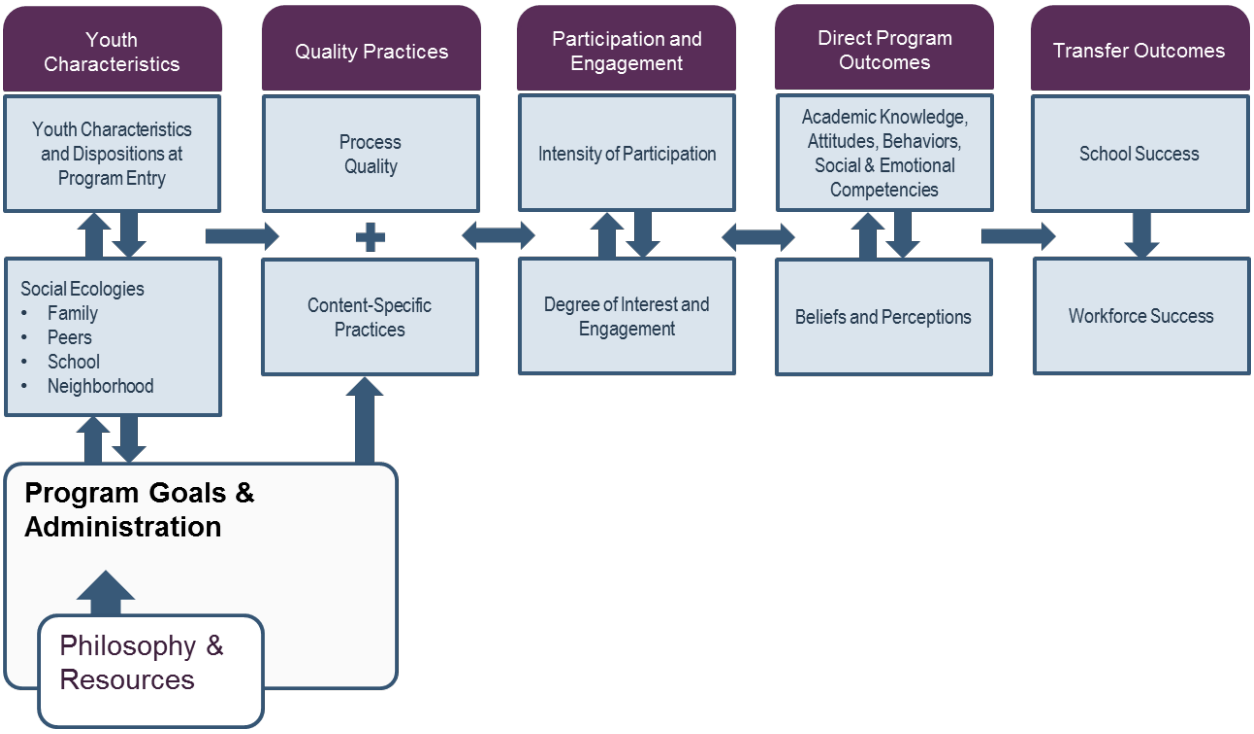
AIR's Conceptual Framework

Before presenting the data, some explanation of AIR's evaluation conceptual framework is warranted. This section, therefore, presents an overview of AIR's understanding regarding how change happens in 21st CCLC programming. This model helps contextualize and situate the data shown in Chapters 3, 4, and 5.

Theory of Change

How does change happen in 21st CCLC programming? That is, from a theoretical standpoint, what elements are important to consider from an evaluation perspective, and how do they fit together? Figure 1 presents the elements AIR considers essential for understanding the 21st CCLC causal story, along with the interrelations of these elements.

Figure 1. How Change Happens in 21st CCLC Programming



Youth Characteristics. To read the figure, begin at the far left, starting with "Youth Characteristics" and "Program Goals & Administration." The framework starts with the youth themselves and how they are influenced and supported by the environments in which they live and go to school. Past programming experiences, relationships with peers and teachers, the level of interest in programming topics and content, expectations regarding program experience, and the level of choice in attending all have a bearing on how youth will engage in and experience 21st CCLC programming (Durlak, Mahoney, Bohnert, & Parente, 2010).

Program Goals and Administration. From the perspective of NJDOE, programs receiving 21st CCLC funding from the state should "supplement the education of students in Grades 4–12 and ... assist students in attaining the skills necessary to meet New Jersey's Core Curriculum Content Standards" (State of New Jersey Department of the Treasury, 2013, p. 1). Overall, New Jersey has defined objectives that outline what is to be achieved in this regard and what this means:

- Goal 1: To provide high-quality educational and enrichment programs that will enable students to improve academic achievement and promote positive behavior and appropriate social interaction with peers and adults

- Goal 2: To implement activities that promote parental involvement and provide opportunities for literacy and related educational development to the families of participating students
- Goal 3: To measure participants' progress and program effectiveness through monitoring and evaluating

NJDOE has defined multiple objectives for each of these goals, which the programs themselves work to implement according to their resources and particular contexts. Each site in New Jersey, therefore, has its own tailor-made goals based on the state's goals but specific to its own population. Each center likewise adjusts its overall administration depending on the student population and the specifics of what it hopes to accomplish. That is, the population to be served, the goals of the program, and the administration of the program (including staffing, professional development, planned activities, recruitment, and the like) will all share a type of dialogical relationship, interacting with one another to give each 21st CCLC program its particular character. The approach of a given 21st CCLC program necessarily takes into consideration its general philosophy (e.g., notions of civic virtue) and the available resources when determining how to structure the program.

Quality. In addition to the predispositions and contextual factors influencing youth before they even enter a program, as well as the program goals and administration, there are various factors that influence the experiences youth have once they are in the program. One such factor is quality; programs must be of high quality to have an impact. Generally, there are two categories of quality: (a) process quality and (b) content-specific practices.

Process quality refers to the adoption of practices and approaches to service delivery that result in the creation of a developmentally appropriate setting for youth, in which participants feel safe and supported and there are opportunities to form meaningful relationships, experience belonging, and be active participants in their own learning and development. These practices are universal because they are truly applicable to any type of youth programming, regardless of content, approach, grade level, or setting.

Content-specific practices are program practices that intentionally cultivate a specific set of skills, beliefs, or knowledge. Often, such practices closely align with the direct outcomes a program is seeking to cultivate in participating youth. For example, content-specific practices include specific approaches to cultivating literacy skills, formal curricula for social and emotional learning, or methods of teaching technology skills. As one might expect, given the broad nature of the overall 21st CCLC program goals and the contextual variance of the specific programs, 21st CCLC program practices vary broadly.

Diversity of Programming. The 21st CCLC program is designed generally to allow for many different types of programming. Beyond the general categories of process quality and content-specific practices, then, what can be said about program quality? Is there a single standard for activity provision—or many standards? Are there aspects of quality that must be evident for a program to have a positive impact on participating youth, regardless of program model or type? These questions are extremely important from an evaluation standpoint.

In relation to the question regarding a single standard, diversity in the design and administration of 21st CCLC programs is supported by the current domain of afterschool research, which suggests that a variety of paths may lead to improved student academic outcomes in both reading and mathematics (Birmingham, Pechman, Russell, & Mielke, 2005; Black, Doolittle, Zhu, Unterman, & Grossman, 2008; Durlak & Weissberg, 2007; Granger, 2008; Lauer et al., 2006; Vandell et al., 2005). That is, in keeping with the overall openness of 21st CCLC, there is not a "one-size-fits-all" formula for program success when it comes to program design and administration.

However, Granger (2008) notes, and important for the purposes of this evaluation, much of the research on the performance of afterschool programs in supporting student growth in terms of academic and behavioral outcomes demonstrates an uneven level of effectiveness. For example, across three especially noteworthy meta-analyses of studies exploring the impact of afterschool programs on student achievement and behavioral outcomes, the majority of the studies in each review did not find evidence that the programs made a significant difference when compared with the outcomes for the control group (Durlak & Weissberg, 2007; Lauer et al., 2006; Zief, Lauver, & Maynard, 2006). Yet, both Durlak and Weissberg (2007) and Lauer et al. (2006) found average positive effects in both academic and nonacademic outcomes, suggesting that a smaller domain of highly effective programs was driving a net average positive effect across programs. Efforts to explore what characteristics, practices, and approaches were especially associated with positive youth outcomes also met with mixed results. Lauer et al. (2006), for example, concluded that various program features seemed to matter at different times but that a consistent pattern across outcomes was lacking.

In contrast, Durlak and Weissberg (2007) had greater success in identifying common characteristics among effective programs. They found that programs focusing on promoting social and personal skills predicated on evidence-based training approaches had significantly positive effects on all the behavioral, attitudinal, and academic outcomes examined in participating youth, with the exception of school attendance. Durlak and Weissberg (2007) concluded that programs should be SAFE—sequenced, active, focused, and explicit—if they are to be effective in meeting a wide domain of behavioral and academic outcomes.

Although this domain of research begins to suggest that it is possible to meet goals and objectives related to the achievement of desired academic and behavioral outcomes by simply paying attention to how programming is delivered (Birmingham et al., 2005; Durlak & Weissberg, 2007), there also is evidence that programs are more apt to accomplish this goal if certain practices are emphasized in intentionally embedded academic-related content delivered to youth at the point of service. For example, Black et al. (2008) demonstrated that students who participated in activities for which there were research-based curricular models and teaching practices designed for the afterschool setting experience statistically significant impacts on mathematics achievement as compared with the outcomes of students enrolled in regular afterschool offerings; however, a similar impact was not found in relation to reading.

Although there has been meaningful progress in uncovering what constitutes quality afterschool programming (e.g., Granger, Durlak, Yohalem, & Reisner, 2007; Little, 2007; Vandell et al., 2005; Wilson-Ahlstrom & Yohalem, 2007; Yohalem, Wilson-Ahlstrom, Fischer, & Shinn, 2009), more recent efforts have predominantly focused on using this knowledge to support the

development of *quality assessment and improvement systems* designed to help afterschool programs better understand (1) what constitutes quality programming, (2) how well the programs measure up to these criteria, and (3) what steps could modify programming to enhance the quality of approaches and offerings. As Granger et al. (2007) note, developing effective quality improvement systems and related interventions is now the most pressing issue facing the afterschool community, including 21st CCLC.

Participation and Engagement. It is not enough that high-quality programming is offered or that youth merely are present for activity sessions. To benefit from a program, students also need to cognitively *engage* in the offerings. One of the most important findings from a study of extended learning time strategies and approaches conducted in South Carolina (Hutchinson, Naftzger, & Miller, 2006) and echoed in a recent report published in relation to AIR's evaluation of the 21st CCLC program in Texas (Naftzger et al., 2012) was that the practices and approaches that support youth engagement in afterschool programming can be quite different depending on the age of the youth in question. For example, because younger elementary students are at the developmental stage at which self-regulatory capacities are beginning to emerge and must be fostered by adults, basic learning is still a novelty and thus can be intrinsically motivating; autonomy is comparatively less important than competence and relatedness. For example, an environment characterized by structured, whole-group activity in which all students are focused on the same task can be completely consistent with the successful facilitation of active student engagement. In contrast, secondary students are at the developmental stage in which autonomy (especially to explore more adult or real-world roles and skills) is an increasingly significant psychological need, and thus adult efforts to control or regulate behaviors can be perceived as an impediment to the satisfaction of this need. As such, for older students to actively engage, there must be opportunities for more choices, more responsibilities, and access to more sophisticated materials and tools. It is important to take such factors into consideration when exploring the relationship between program quality and youth engagement at the point of service.

***Direct Program Outcomes*.** If youth engage in quality activities over multiple sessions, they are likely to change in ways that are a *direct consequence* of 21st CCLC participation. These more immediate *direct program outcomes* can fall within a wide spectrum of categories, including social-emotional learning, critical thinking and decision making, and initiative and self-direction (Wilson-Ahlstrom, Yohalem, DuBois, & Ji, 2011). In addition, they youth are likely to acquire specific, content-related skills in areas such as reading and mathematics, particularly if there has been an effort to build specific skills in individual students.

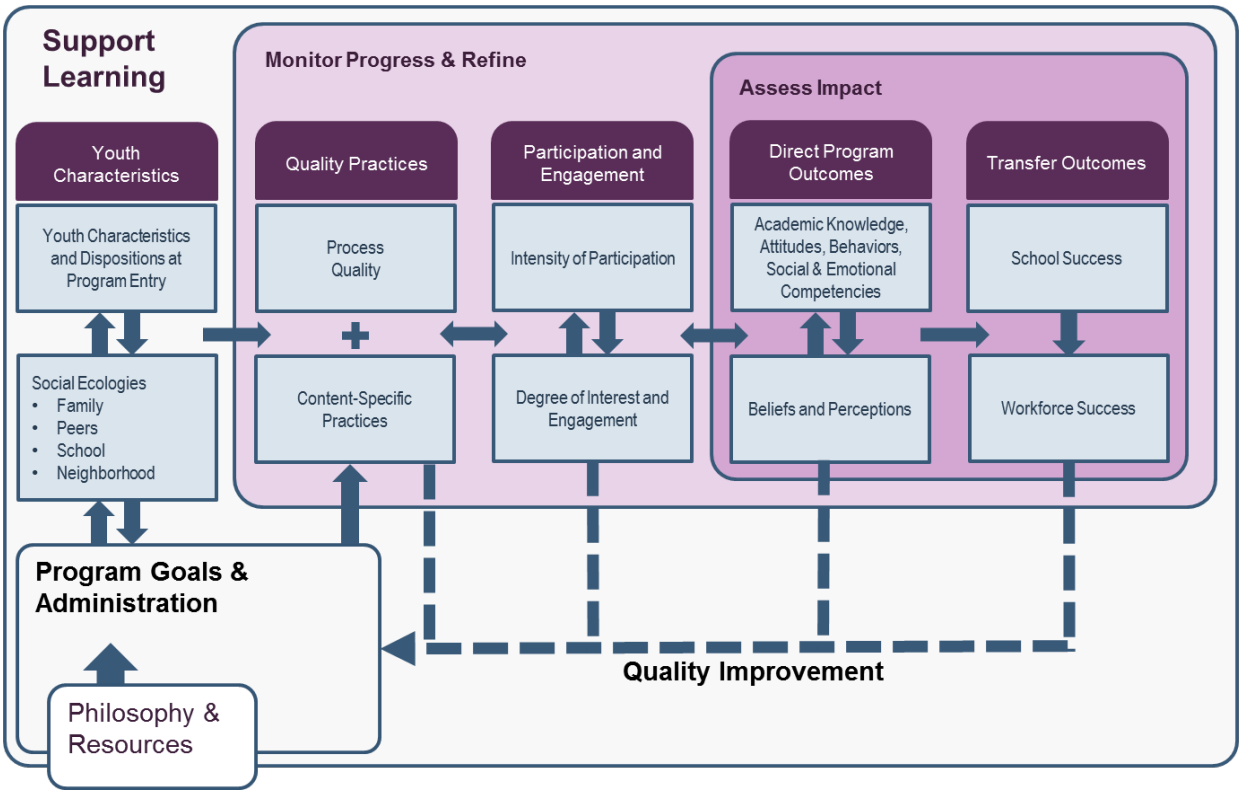
***Transfer Outcomes*.** Eventually, it is likely that through the change process described earlier (in which youth participation in high-quality programming yields desired direct program outcomes in keeping with well-structured program goals), youth will have greater school success and, in turn, greater workforce success than they would have had without the program. Some aspects of school success, such as chronic absenteeism and disciplinary rates, as well as grades and assessment scores (which are of great interest to both NJDOE and the U.S. Department of Education), can be and have been assessed using evaluation efforts. These outcomes are slightly longer term than those noted as direct program outcomes and are theorized to be evident after a year or so of regular participation. More broadly, transfer outcomes such as long-term school or workforce success remain somewhat underexplored domains of research. Simply stated, it is much more difficult to analyze these types of outcomes given the longer time periods involved

and the obstacles to controlling for confounding variables. Nevertheless, the 21st CCLC program's eventual goal is to improve these longer term transfer outcomes, and this goal is an important part of the conceptual theory of change.

AIR and the Theory of Change

Figure 2 presents the same theory of change as Figure 1 but with added components showing the role of an evaluator.

Figure 2. Evaluation and the Theory of Change



Four additional elements bear explanation, and each of them reflects a different type of evaluation-related work. First, taking into consideration all components of the theory of change, the evaluator should explore *impact*. That is, what effect does participation in the 21st CCLC program actually have on participating youth? This is a causal question and is usually the component that most outside observers think of with respect to evaluation's role. However, as Figure 2 shows, *assess impact* is focused on the far-right two columns in the theory of change, and their components relate to outcomes.

More broadly—and with respect to program goals, program implementation, youth participation, and outcomes data—evaluation includes *monitoring progress and refining*. That is, it is especially important to view data from the four right-hand columns to assess program performance relative to goals. These types of data can be presented as basic descriptive statistics or, in response to the U.S. Department of Education's new focus as part of the latest Elementary and Secondary Education Act (ESEA) authorization, as program or performance indicators.

Programs themselves or NJDOE can use these types of data to determine valuable questions for subsequent evaluation efforts, to identify areas of program weakness or need, and to assess whether earlier improvement efforts are on track. It is important to note, however, that data associated with monitoring progress and refining are suggestive and are not presented as program effects—even if the data in question are outcomes data.

Similarly, none of the data analyses implied by assess impact and monitor progress and refine are of much use if they are not used to improve programming. That is, both indicator-type data and impact analyses should be folded into quality improvement efforts. This can be done at both statewide and program levels, though these efforts will look different given their scale. At the state level, such efforts likely drive refinements to requests for proposals (RFPs), grantee training opportunities, instructional materials, and similar guidance, while program-level improvement efforts should be more tailor-made and intentional in their logic. In this latter respect, however, the state may helpfully provide structure for grantees to use their data via quality improvement systems.

Finally, with respect to the entire conceptual framework, the evaluator should enable and support learning. This should be interpreted broadly. For instance, with respect to outcomes, interpretation of evaluation results in user-friendly terms is important; for monitoring progress and refining, presenting data in a way that is useful for reflection and project planning relative to goals may be most important; and providing structure around quality improvement may be critical for ensuring data analyses are incorporated into a continuous improvement process. There are many aspects to this, which typically evolve over the course of a multiyear evaluation process. However, AIR considers this to be an important element of evaluation work.

Chapter 3. Grantee, Center, and Student Characteristics

Programs funded by 21st CCLC grants are often characterized by a wide diversity of approaches, student populations, and types of organizations involved in providing 21st CCLC programming. This chapter summarizes the characteristics of grantees, centers, and students associated with 21st CCLC programs active during the 2015–16 school year. Overall, there were 52 grantees (compared with 50 grantees during 2014–15) operating 130 centers (compared with 124 centers during 2014–15). In all, these 130 centers served 15,449 youth (compared with 15,049 during 2014–15).[4]

Grantee Characteristics

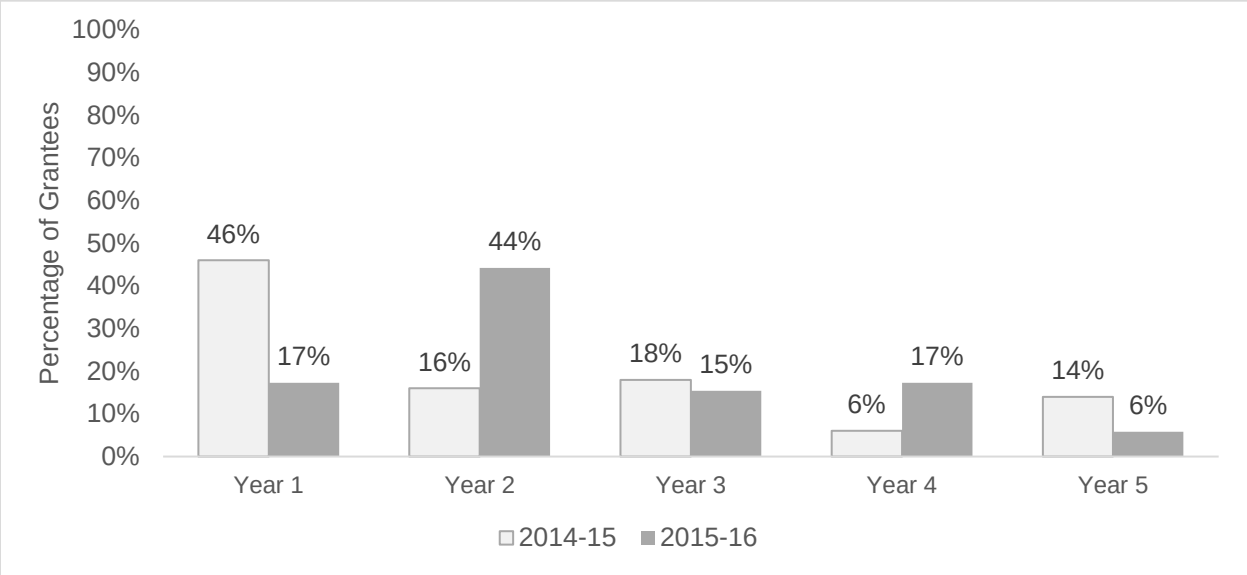
This section contains information on key grantee characteristics. In this report, the term *grantee* refers to the organization that serves as the fiduciary agent on the grant in question, whether it is a school district, community-based organization, or other entity and whether it is ultimately responsible for administering grant funds at the program level.

Grantee Maturity

Programs evolve across the grant period. For example, grantees may find themselves needing to emphasize some elements of their programs and reducing or eliminating others in response to changes in the students served. In addition, it would be optimal for grantees, over time, to be learning how to (1) provide more effective and engaging programming for youth and (2) more meaningfully embed academic content in their program offerings in ways that address the needs of the students they are serving. As Figure 3 shows, the majority of the grants active during the 2015–16 school year were in Year 2 of funding (not surprising, given that in 2014–15 the majority were in Year 1). New Jersey's 21st CCLC grants are for five years, so many of the programs active during Year 2 could still be considered new, though perhaps more experienced, focusing their efforts on fine-tuning their programs after a single year of operation.

[4] Note that the number of sites and centers include those that may not have operated the full year. Only 123 centers provided school-year attendance data.

Figure 3. Number of Grantees by Year of Operation, 2014–15 and 2015–16

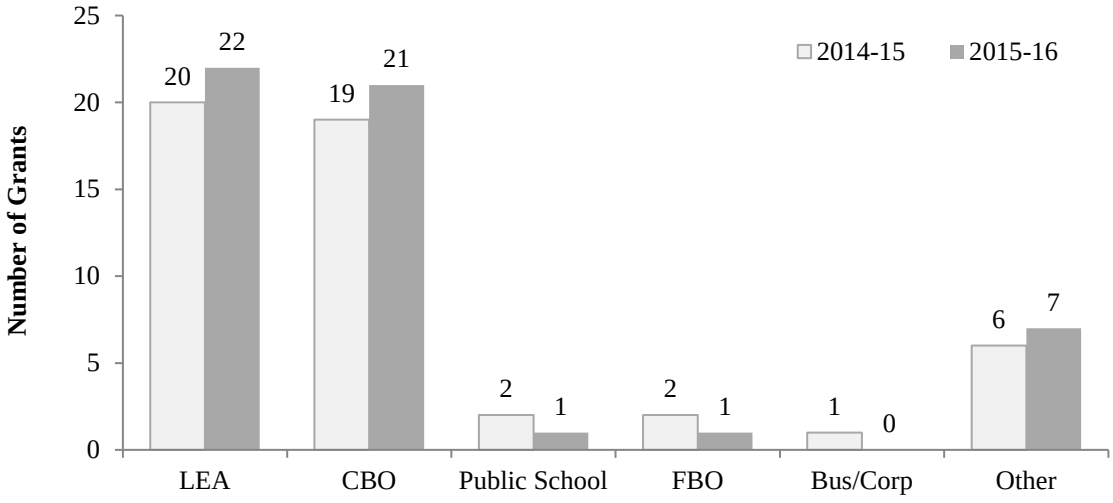


Source. PARS21.

Grantee Organization Type

An important element of the 21st CCLC program is that all types of organizations are eligible to apply for and receive 21st CCLC grants. As Figure 4 shows, 43% of grants active during the 2015–16 school year were held by school districts (up slightly from 40% the previous year), while community-based organizations accounted for 40% of the grants active during this period (up from 38% the previous year). Public schools and faith-based organizations each accounted for only about 2% (compared to about 4% last year), while all other categories accounted for roughly 14%. Grant types therefore remained about the same between 2014-15 and 2015-16.

Figure 4. Number of Grantees by Organization Type



Source. PARS21.

Key Center Characteristics

This section presents key center-characteristic data. In this report, the term *center* refers to the physical location where 21st CCLC–funded services and activities take place. Each center has defined hours of operation, dedicated staff members, and a site coordinator to manage operations. Each 21st CCLC grantee in New Jersey has at least one center; many grantees have more than one center.

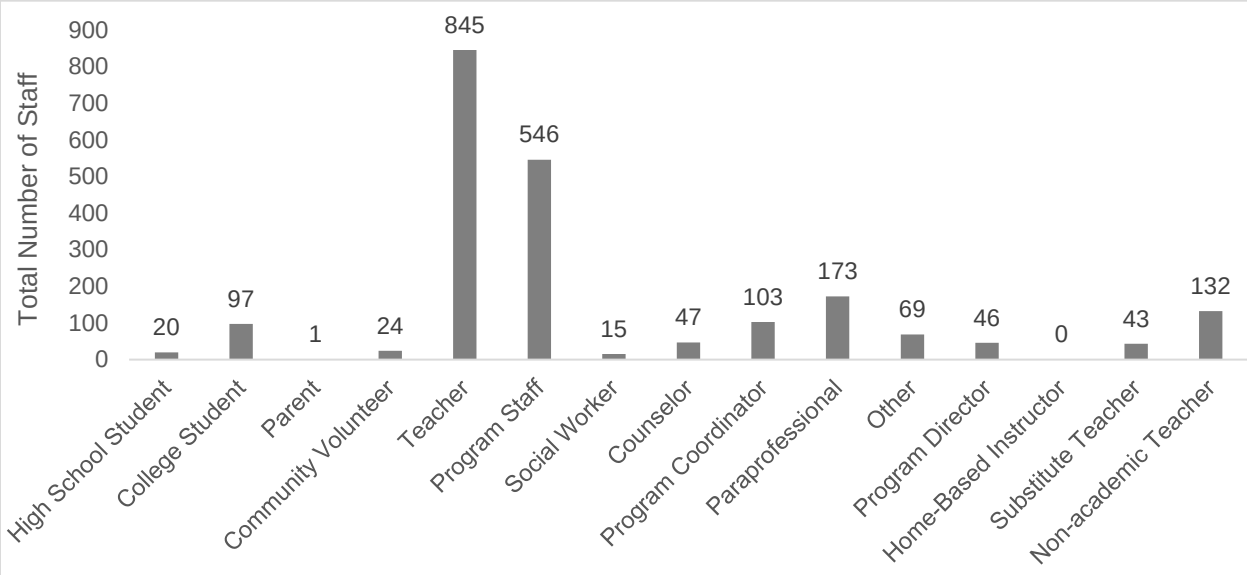
Center characteristics can be described either as indicative of research-supported best practices or as innate attributes of the center in question without a strong connection to the afterschool quality practice literature. Center characteristics indicative of the latter might include the grade level served, program maturity, and organizational type. For example, identifying a program as one that serves only elementary students says nothing about the quality of that program.

Other characteristics of a site, such as the staffing model, are still somewhat ambiguous when viewed from a quality practice standpoint, with the literature unclear on the superiority of certain staffing approaches. From a policy standpoint, NJDOE considers certain approaches to staffing for certain types of activities to be appropriate from a quality standpoint—namely, that certified teachers should staff academic programming provided in the afterschool program.

Staffing

Grantees in New Jersey report staff information in PARS21, linking each staff member to activities provided during 21st CCLC programming. Staff can be categorized in a number of different ways, such as “parent” and “college student.” Counting only those staff that were in some way associated with the provision of actual activities, a 2,161 staff were reported by grantees for school year 2015–16 across all programs. In terms of classification of these staff, by far the most commonly reported staff types were “teacher” (39.1% of all staff) and “program staff” (25.3% of all staff), with a distant third being “paraprofessional” (8.0%) followed by “nonacademic teacher” (6.1%) and “program coordinator” (4.8%). Figure 5 shows the total number of staff across New Jersey by staff type.

Figure 5. Total Number of Staff by Staff Type



Note. Based on activity staff data for 123 centers.

Overall, centers had an average of 17.6 total staff for the school year, with a median of 16 (again, only counting staff who actually participated in activity offerings). However, as Table 1 shows, there was some variation in total staff, with a standard deviation of 10.7 staff members.

Table 1. Overall Statistics on Number of Center Staff

	N	Mean	Median	Minimum	Maximum	Standard Deviation
Total Staff	123	17.6	16	3	64	10.7

Looking at individual staff categories, all median values were zero except for teachers (median four), program staff (median two), and program coordinators (median one). On average, programs had about seven teachers and four program staff, along with one or two other types of support staff (e.g., a coordinator). A total of 42 programs (34.1%) reported relying very heavily on teachers, with 50% or more of school-year activity providers being teachers. For comparison, only 28 programs (22.8%) reported that 50% or more of their activity staff were program staff. A small number of programs, five total (4.1%), reported a heavy reliance on college students, with more than 50% of their staff as college students.

In addition to exploring the number of staff employed by centers during the 2015–16 school year, researchers calculated the average student-to-staff ratio associated with activity sessions provided during the span of the school year in question. As Table 2 shows, the average student-to-staff ratio was approximately one staff member for every 11 or so youth participating in activities, although across centers, the span of ratios was quite broad, ranging from just under four students to approximately 36. These ratios, however, did not change much between 2014–15 and 2015–16; the table shows both years for ease of comparison.

Table 2. Average Student–Teacher Ratio per Center, 2015–16

	N	Minimum	Maximum	Mean	Standard Deviation
2015–16 student–staff ratio	123	3.88	36.49	11.14	5.11
2014–15 student–staff ratio	116	1.81	32.64	11.49	5.34

Source. PARS21.

Grade Levels Served

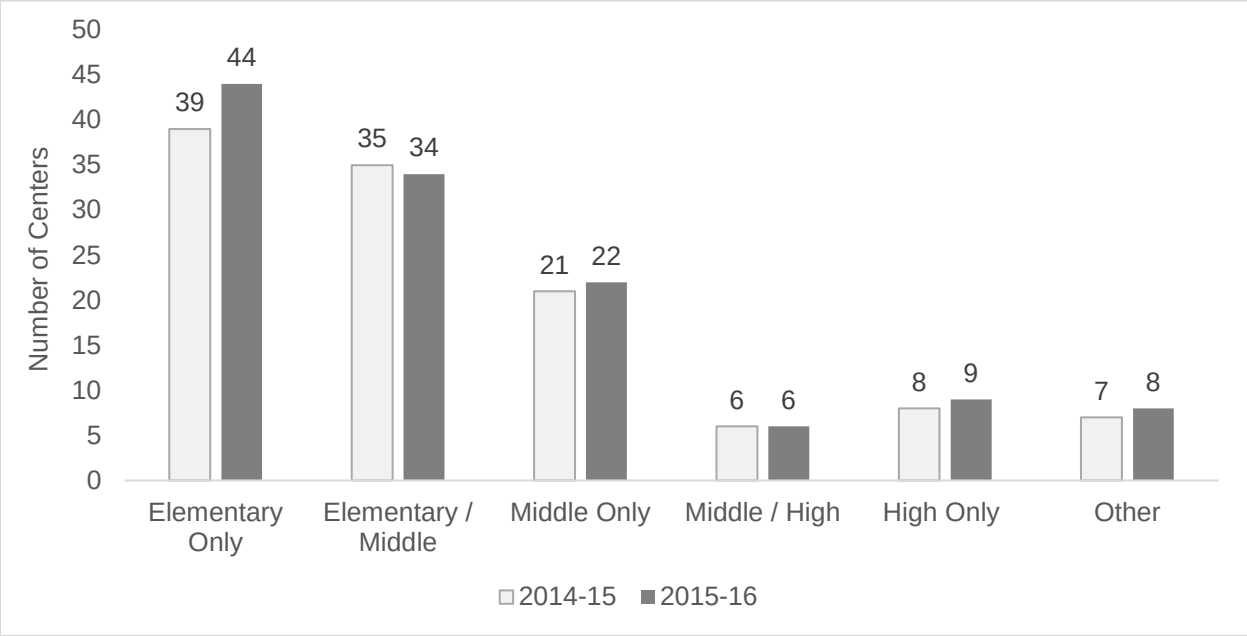
A topic garnering increasing attention on the federal stage relates to the role grade level plays in (1) how 21st CCLC programs should structure their operations and program offerings and (2) the domain of outcomes they should be accountable for through performance indicator systems. Using student-level data on the grade levels of students attending centers, centers active during the 2015–16 school year were classified as follows:

- *Elementary Only*, serving students up to Grade 6
- *Elementary/Middle*, serving students up to Grade 8
- *Middle Only*, serving students in Grades 5–8
- *Middle/High*, serving students in Grades 5–12
- *High Only*, serving students in Grades 9–12

A sixth category, called *Other*, includes centers that do not fit one of the five categories and includes centers that serve students across all three grade levels or some other combination of grade levels.

The *High Only* category is especially important to analyze because afterschool programming for older students often looks considerably different from programming for elementary or middle school students (Naftzger et al., 2007). In addition, high school students have different needs from younger students, and they often have other afternoon obligations, such as jobs or extracurricular activities. The bulk of the centers active during the 2015–16 school year served elementary or middle school students in some capacity (constituting 82.1% of all sites), whereas not quite two thirds of all sites served elementary students in some capacity (60.2% of all sites). These figures are not substantially different from those observed the previous year, as Figure 6 shows (note the slightly lower overall center count for 2014–15, however).

Figure 6. Number of Centers by Grade Level Served



Note. Based on 116 centers for 2014–15 and 123 centers for 2015–16.

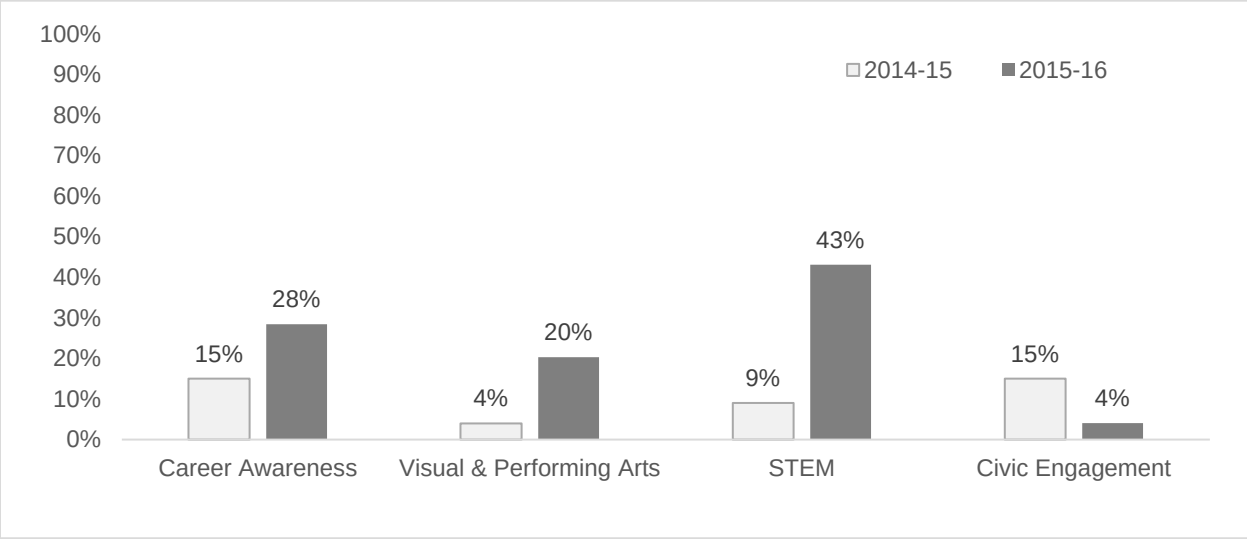
Activity Themes

For the 2015–16 school year, grantees were required to adopt one or more themes when providing activities. The grantees were to select a theme based on the students' needs, interests, and developmental age and were meant to further support targeted skill building and development through the provision of activities youth would especially find engaging. Themes included the following:

- Science, technology, engineering, and mathematics (STEM);
- Career awareness and exploration;
- Civic engagement; and,
- Visual and performing arts.

Prior to 2014–15, grantees were not required to select a theme. Over time, however, it is likely that the percentage of sites offering activities consistent with a theme will increase as older grants that do not have a theme become inactive. This is, in fact, what seems to be taking place, as 95.9% of centers active during the 2015–16 school year have provided activity sessions associated with one or more of the aforementioned themes, based on data reported in PARS21 (compared with about 42% of centers the previous year). As Figure 7 shows, 28% of centers reported a career awareness theme, 20% visual and performing arts, 43% STEM, and 4% civic engagement. Themes were derived for centers based on (a) whether they offered any activities associated with a given theme and (b) how many total activity minutes were associated with each theme the center reported (with the theme designation going to the theme that had the highest minutes).

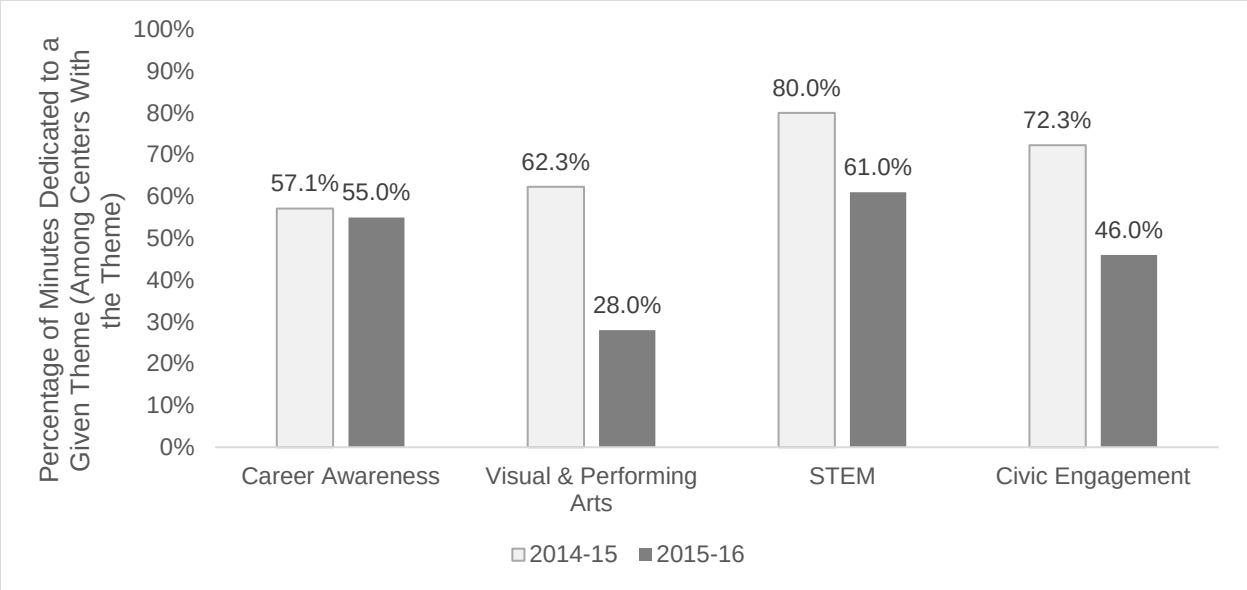
Figure 7. Percentage of Centers Offering Activities Linked to a Given Theme



Source. PARS21.

As Figure 8 shows, centers with a career awareness theme spent, on average, about 55% of their total activity minutes on career awareness. Centers with a visual and performing arts theme spent 28% on such activities. Centers focusing on STEM spent about 61% of their time on such activities, and centers with a civic engagement focus spent about 46% of their time on the theme. In each case, the percentage of time dedicated to each theme dropped year to year. This could have resulted simply from an increase in activities generally (as reported in PARS21) or to a difference in reporting. However, basic programmatic changes also are possible and should be considered as an explanation.

Figure 8. Percentage of Total Activity Minutes Dedicated to Activity Themes, Among Centers With Each Theme



Source. PARS21.

Student Characteristics

During the 2015–16 school year, 15,449 students participated at some level (i.e., attended programming for at least one day during the school year) in 21st CCLC programming at 123 active centers for which the researchers had data during this period.[5] This population was diverse in terms of ethnicity, gender, grade level, and economic level, as Table 3 shows. Generally, the population of students served during the 2015–16 school year was Black and Hispanic/Latino; was enrolled in elementary or middle school, especially in Grades 4–6; and was eligible for the free or reduced-price lunch programs. In terms of year-to-year changes, there was a slight drop in the proportion of white students and a slight increase in the proportion of black students. Other statistics remained moderately stable year to year.

Table 3. Summary of Demographic Information for Students, 2015–16

	Demographic Category	2014–15		2015–16	
		Number of Students	Percentage	Number of Students	Percentage
Race/ Ethnicity	White	2,199	14.6%	1,713	11.1%
	Black	4,820	32.0%	5,393	34.9%
	Hispanic/Latino	6,592	43.8%	6,936	44.9%
	Asian	386	2.6%	343	2.2%
	Native American	44	0.3%	28	0.2%
	Pacific Islander	26	0.2%	20	0.1%
	Unknown	982	6.5%	1,016	6.6%
Gender	Male	7,705	51.2%	7,876	51.0%
	Female	7,344	48.8%	7,573	49.0%
Grade Level	4	2,853	19.0%	3,051	19.7%
	5	2,489	16.5%	2,664	17.2%
	6	2,501	16.6%	2,739	17.7%
	7	1,940	12.9%	1,960	12.7%
	8	1,821	12.1%	1,661	10.8%
	9	907	6.0%	1,097	7.1%
	10	735	4.9%	702	4.5%
	11	627	4.2%	554	3.6%
	12	331	2.2%	252	1.6%
Free or Reduced-Price Lunch	Reduced	1,105	7.3%	1,261	8.2%
	Free	10,003	66.5%	10,347	67.0%
	Not Available	3,941	26.2%	3,841	24.9%

Source. PARS21.

[5] One hundred sixteen centers active during the 2014–15 school year had student-level attendance records in PARS21, confirming participation in actual activity sessions during the span of the school year.

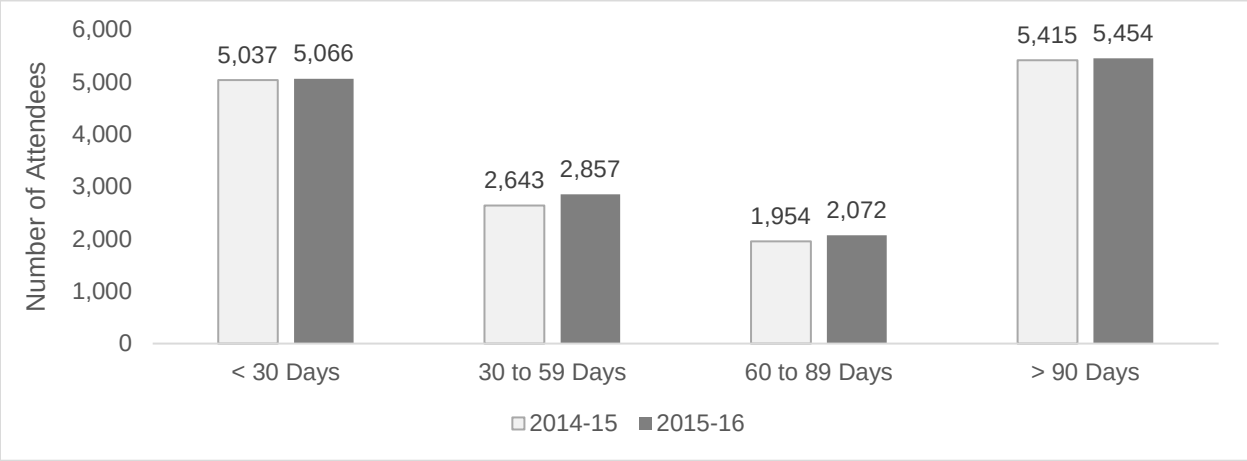
Student Attendance Levels

Attendance is an intermediate outcome indicator that reflects the potential breadth and depth of exposure to afterschool programming. In this regard, attendance can be considered in terms of the (1) total number of students who participated in the center's programming throughout the course of the year and (2) frequency and intensity with which students attended programming when it was offered. The former number can be used as a measure of the breadth of a center's reach, whereas the latter can be construed as a measure of how successful the center was in retaining students in center-provided services and activities.

Among students participating in activities during the 2015–16 school year, the average number of days attending 21st CCLC programming was 66—the same as the previous year. Figure 9 shows the student population served during the 2015–16 school year broken into four attendance gradations: the percentage of students attending fewer than 30 days, students attending 30 to 59 days, students attending 60 to 89 days, and students attending 90 or more days. As Figure 9 shows, one third of the students (32.8%, compared with 33.5% the previous year) attended fewer than 30 days, a level consistent with previous years, and slightly more than one third participated for 90 or more days (35.3%, compared with 40% the previous year). These attendance levels are fairly consistent with previous year attendance levels.

To demonstrate program impact, one would hope that there would be a positive relationship between higher levels of attendance in the program and the likelihood of gains in student achievement and behavioral outcomes. For this reason, attendance rate is incorporated into the impact models presented in Chapter 5.

Figure 9. Number of Students Served in 21st CCLC by Attendance Gradation



Source. PARS21.

In addition to levels of program attendance during the 2015–16 school year, the research team was interested in exploring the extent to which students participating during this period had been attending the program at a given center beyond the school year in question. Hypothetically, it would be expected that a higher number of years of continuous participation in the program would be associated with a greater degree of improvement on the outcomes of interest in this report. However, as Table 4 shows, for the vast majority of students (nearly 70%), the 2015–16

school year represented the first year in which they participated in 21st CCLC programming at the center in question; approximately 22% were in their second year of participation during the 2015–16 school year. Three or more years of continuous participation was found to be relatively rare. The results were very similar to the same analysis conducted with the previous year's data.

Table 4. Continuous Years of Student Participation, 2015–16

	2014–15		2015–16	
	Number of Students	Percentage	Number of Students	Percentage
1 year	10,468	69.7%	10,706	69.3%
2 years	3,131	20.8%	3,385	21.9%
3 years	1,083	7.2%	906	5.9%
4 years	289	1.9%	355	2.3%
5 years	49	0.3%	98	0.6%
6 years	2	0.0%	2	0.0%
7 years	0	0.0%	0	0.0%
8 years	1	0.0%	0	0.0%

Note. Prior-year records were matched to 15,452 students using participant identifiers. One year of continuous participation, for example, indicates that a given student is either in his or her first year of programming during the 2015–16 school year or that there was an interruption in participation prior to the 2015–16 school year.
Source. PARS21.

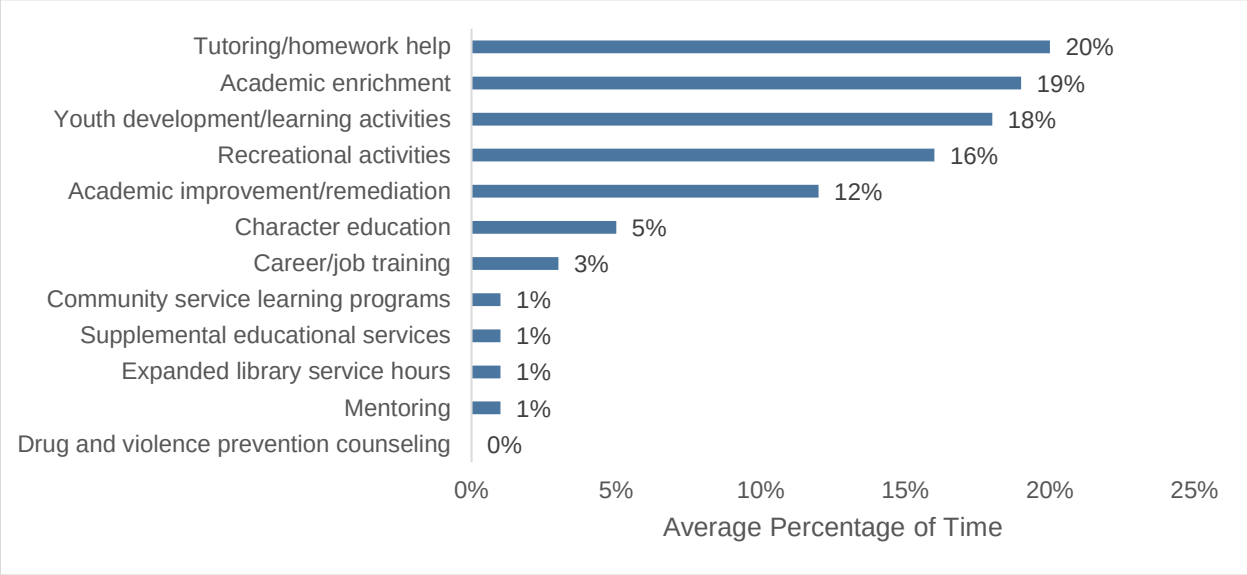
Student Attendance by Activity Types

An effort was made to determine how much time 21st CCLC participants spent in activities of different types. Within PARS21, activities in which attendees participate can be classified according to the following different types:

1. Academic improvement/remediation;
2. Academic enrichment;
3. Tutoring/homework help;
4. Mentoring;
5. Drug and violence prevention counseling;
6. Expanded library service hours;
7. Recreational activities;
8. Career/job training;
9. Supplemental educational services;
10. Community service learning programs;
11. Character education; and,
12. Youth development/learning activities.

Using these activity categories, participant attendance records, and activity session duration data, a total number of minutes for each activity type was calculated for each participant. This information was then used in conjunction with total participation minutes to derive student-level percentage statistics concerning each attendee's time spent in each type of activity. Averages of these percentages were then taken to determine, on average, how much time participants spent in each activity category. Figure 10 shows the results.

Figure 10. Proportion of Time Each Participant Spends on Activities of a Given Type (Averages)



General statistics were also run for total participant hours (school year) by activity type, calculating the average and median number of total hours for each type of activity alongside a minimum, maximum, and standard deviation (see Table 5). Tutoring, again, was highest in terms of the average number of total hours, with 35 school-year hours. However, the median value for tutoring was only 2 hours (with a high standard deviation of 51), indicating that this average was driven by a smaller number of youth with very high hour allotments for tutoring. In terms of median values, academic enrichment had the most hours (12), followed closely by youth development/learning activities (9) and recreational activities (8). These median values indicate that a larger proportion of youth participated at higher levels in enrichment, youth development, and recreation activities than in tutoring.

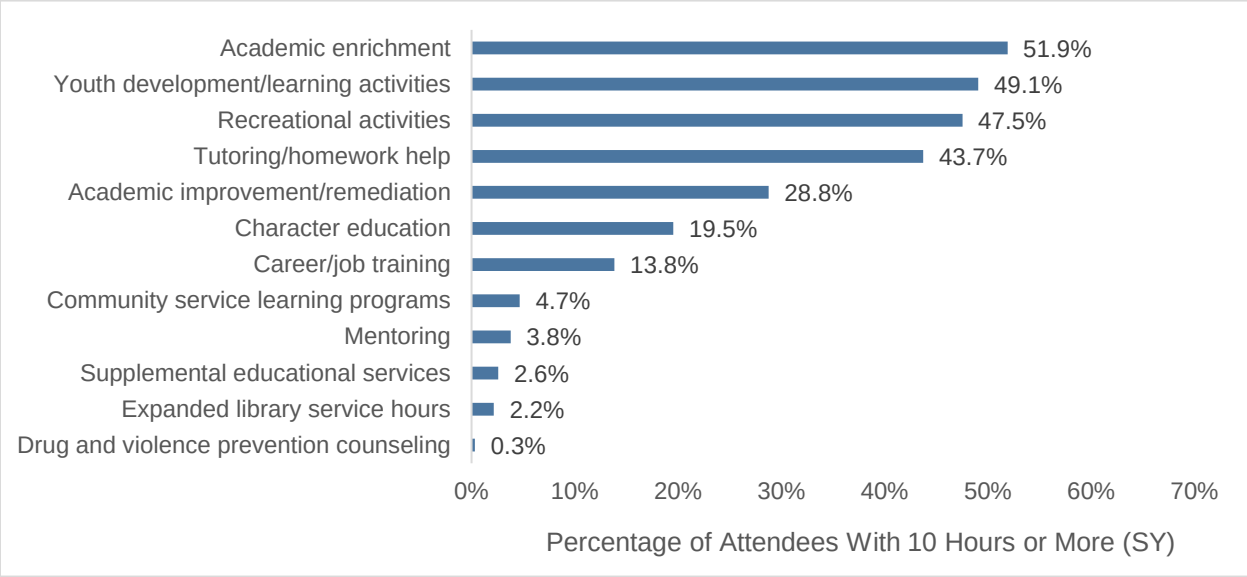
Table 5. Total School-Year Hours of Attendee Participation, by Activity Type

	Mean	Minimum	Maximum	Median	SD
Academic improvement/remediation	16	0	277	0	34
Academic enrichment	29	0	241	12	38
Tutoring/homework help	35	0	333	2	51
Mentoring	2	0	148	0	11
Drug and violence prevention counseling	0	0	26	0	1
Expanded library service hours	1	0	84	0	5
Recreational activities	23	0	331	8	36

	Mean	Minimum	Maximum	Median	SD
Career/job training	5	0	281	0	14
Supplemental educational services	1	0	314	0	15
Community service learning programs	1	0	56	0	5
Character education	7	0	292	0	18
Youth development/learning activities	24	0	225	9	34

To explore the intensity of youth participation in each activity category type, a simple calculation was made to identify youth participating in at least 10 hours in each activity type (again, counting total hours for the entire school year). Figure 11 shows the percentage of youth participating for at least 10 hours. As indicated, academic enrichment was the highest, with nearly 52% of all youth participating for 10 hours or more during the year, followed closely by youth development/learning activities (49%) and recreation (about 48%). Approximately 44% of youth participated in tutoring activities for 10 hours or more.

Figure 11. Percentage of Attendees With 10 or More Hours in a Given Activity Type (School Year)



Participation in Reading and Mathematics Activities

Another approach to examining students' participation in 21st CCLC programming offered during the 2015–16 reporting period is to explore the extent to which students participated in activities meant to support skill building in mathematics and reading, regardless of activity type (e.g., enrichment, tutoring, academic remediation). As mentioned earlier, one of the central goals of the 21st CCLC program is to support student growth and development in reading and mathematics. As Table 6 outlines, students on average participated in approximately 54 hours of reading/literacy programming during the 2015–16 reporting period and 50 hours of mathematics programming. In comparison with 2014–15, these hour averages are far higher and might seem like a vast improvement; however, the leap between years almost certainly results from a change in the way these data were reported in PARS21. In 2014–15 and earlier, activity records in PARS21 were reported as targeting a single subject variable, along with an activity duration. A

simple calculation thus provided total subject hours. In 2015–16 and later, activity records were allowed to indicate multiple subjects as a target (appropriate for, say, an enrichment activity that might equally incorporate elements of mathematics, science, and reading). Because of this recording difference, many more activities' minutes were counted toward each subject hour total than they were in previous years. Comparing years to each other is, therefore, like comparing apples and oranges; however, the difference between years does show what a difference a data collection method can make.

Table 6. Average Number of Hours in Reading and Mathematics per Student, 2015–16

	N	Minimum	Maximum	Mean	Standard Deviation
2015–16 reading/literacy education activities	15,452	0	671*	54.0	68.3
2015–16 mathematics education activities	15,452	0	661*	49.5	66.0
2014–15 reading/literacy education activities	15,049	0.75	101	29.16	25.23
2014–15 mathematics education activities	15,049	0.25	54	25.68	19.68

Source. PARS21. Note that the method of activity data reporting changed in 2015–16 to allow for activity records to target multiple subjects.
* These values are fairly extreme outliers; only 15 students had more than 400 hours total of either mathematics or reading. For more context, the median mathematics and reading values are 21 hours and 27 hours, respectively (for 2015–16).

Chapter 4. Staff Survey Data

According to the 21st CCLC grant coordinators, all front-line 21st CCLC staff receive a staff survey every year during November and December. The survey is provided through an online module (ETRS), with coordinators inputting a list of all staff who should take the survey at each site. Coordinators are instructed to include only staff who fit the following description:

- Need to spend the majority of their time working directly with youth by leading activities;
- Can be paid or are volunteers;
- Can work for the grantee directly or for a partner organization;
- Are working in the program or recently ended their participation in the program (e.g., led activities in a previous session that was active at some point in 2015); and,
- Should work (or did work) in the program at least once a week.

The staff members whom the coordinator enters into the system are then sent a link to the survey; the coordinator is able to monitor completion by staff member.

The staff survey asks the identified 21st CCLC front-line staff about their perceptions of program practices and procedures at their particular site. Specifically, the questions relate to the following areas:

- Program goals and vision;
- Perception of collective staff efficacy;
- Youth autonomy and choice;
- Overall program design;
- Communication and linkages to the school day;
- Data use;
- Service delivery practices;
- Youth ownership;
- Internal communication; and,
- Parent communication.

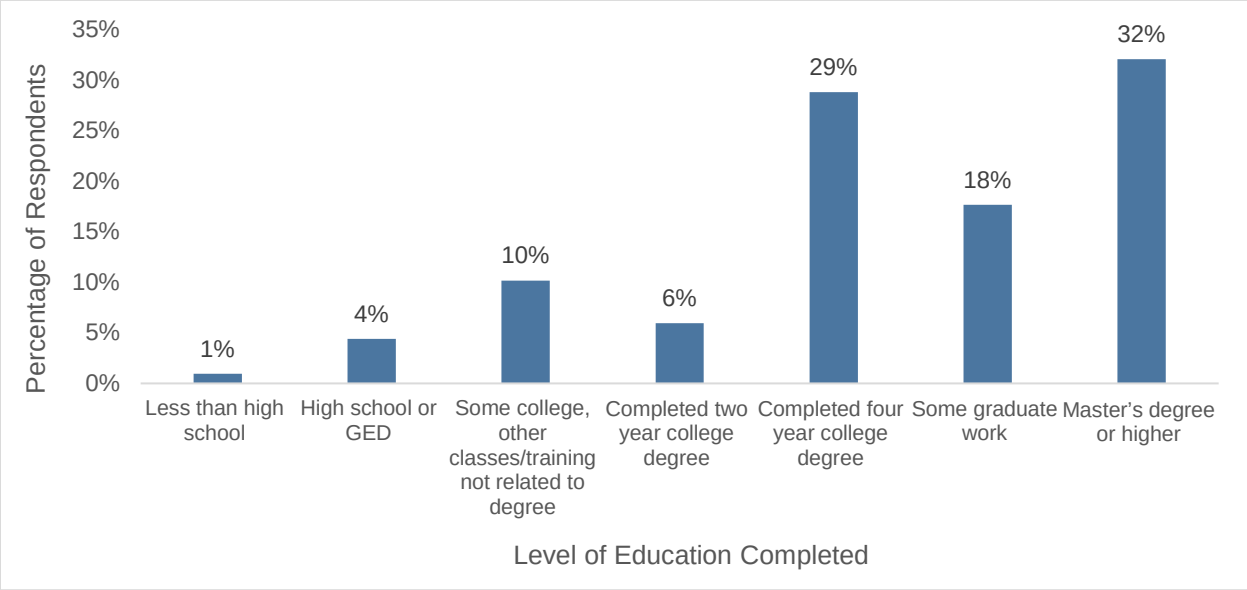
Each of these areas includes a bank of questions intended to work together as a single scale. In previous years, these questions were consequently analyzed using Rasch analysis to create scale scores for each area, with scores converted into center-level scale scores for presentation in the 21st CCLC leading indicators. For 2015–16, however, as part of the overall data-collection transition and as a way to free resources for transition-related development tasks, the leading indicators were not calculated. For 2015–16, then, the staff survey data are presented as simple overall response rates to each individual question.

Staff Survey Response Rates and Respondent Population Note

Overall, 546 responses were received in December 2015 and January 2016. Of these, 530 were completed entirely. These surveys were received from 67 centers associated with 30 grantees, with an average of 8.1 and a median of 8.0 surveys per center. Many of the newest grantees were not included in the staff survey (per NJDOE guidance); nor were grantees included that had just concluded programming. Therefore, this number, despite being significantly lower than the overall grantee count cited in Chapter 3, does not indicate a low response rate. In fact, only two grantees, representing six centers, out of those required to respond did not respond. The completion rate, then, was around 92% of the grants NJDOE indicated should receive the survey. This means, therefore, that the survey results shown in this report are not necessarily indicative of all grantee staff opinions; rather, these responses reflect opinions of staff who worked at grants that were not considered "new" by NJDOE.

The survey asked a few questions about each respondent to gain a sense of the staff serving 21st CCLC programs. On average, those who responded to the survey reported spending about 10.9 hours per week in their program, on average (median of 9.0), and working with an average of 29 students per day (with a median of 20).[6] Most staff reported having a teaching credential or certification (69.2% of 522 responses provided), and almost four out of every five staff reported having at least a four-year degree (79%, with about a third of all staff; 32%, reporting a master's degree or higher). See Figure 12 for all responses regarding education level.

Figure 12. Level of Education Completed by 21st CCLC Staff Responding to the Staff Survey



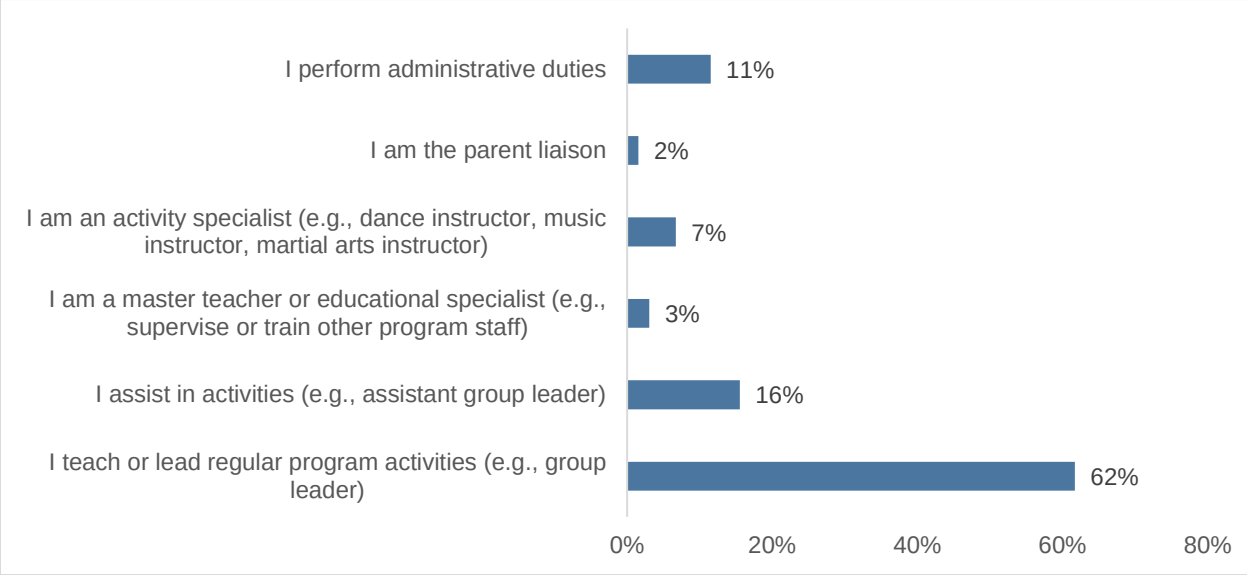
Note. Based on 521 answers to the question, "What is your highest level of education?"

Survey respondents were also asked to classify their role relative to the program. Respondents were presented with six broad categories and were asked to select the one that *best* describes their *primary* role in the program. Slightly fewer than two thirds of all respondents selected "I

[6] Information about hours per week and the number of students worked with each day is based on 524 received responses.

teach or lead regular program activities," as might be expected, with about 16% indicating they primarily *assist* in activities and 11% selecting *administrative duties* as their primary role. See Figure 13 for a complete presentation of all responses to this question.

Figure 13. Staff Roles in 21st CCLC Programming Among Survey Respondents



Note. Based on 522 responses to the question "Which of the following best describes your primary role in the program?"

All of the characteristics in Figure 13 belong to survey respondents and do not necessarily characterize all 21st CCLC staff.

Survey Responses by Survey Area

Respondents were given a series of general program goals and asked to indicate the extent to which that particular kind of goal applies to their particular program. Responses provided were "Not a goal," "Secondary goal," and "Primary goal." By far the most common primary goal was "provide students with access to academic enrichment opportunities," followed by "raise the academic performance levels of *any* students who have an interest in participating." As Table 7 shows, most programs indicated that the goals listed were either a secondary or primary goal.

Table 7. Responses Concerning Program Goals and Vision

What, in your view, are the goals of the afterschool program?	Not a Goal	Secondary Goal	Primary Goal	N Responses
a. Enable the *lowest-performing* students to achieve grade-level proficiency	7.2%	27.5%	65.3%	539
b. Raise the academic performance levels of *any* students who have an interest in participating	2.0%	13.8%	84.2%	543
c. Provide supervised space for students to complete homework	3.7%	24.0%	72.3%	541
d. Provide opportunities for students to participate in activities not offered during the school day	2.0%	18.2%	79.7%	538

What, in your view, are the goals of the afterschool program?	Not a Goal	Secondary Goal	Primary Goal	N Responses
e. Provide students with access to academic enrichment opportunities	0.7%	9.4%	89.9%	542
f. Enhance the social or civic development of students	1.7%	22.9%	75.4%	533
g. Enhance the artistic development of students (e.g., visual and performing arts)	3.6%	38.5%	58.0%	533
h. Provide students with the opportunity to participate in sports and recreation activities	5.4%	40.4%	54.2%	539
i. Other	46.2%	12.9%	40.9%	93

Respondents were then asked to indicate how much they agree with a series of statements relating to staff efficacy. As Table 8 shows, about 90% of respondents agreed (indicating either "agree" or "strongly agree") with each specific staff efficacy practice described. The choices "actively and continuously consult and involve youth" and "provide structured and planned activities explicitly designed to help youth to get to know one another" received the highest levels of agreement.

Table 8. Responses Concerning Collective Staff Efficacy

Please rate the extent to which you agree or disagree with the following statements regarding all staff that work with students in this program:	Strongly Disagree	Disagree	Agree	Strongly Agree	Not Sure	N Responses
a. Program staff listen to youth more than talk at them.	1.3%	5.0%	49.4%	41.2%	3.2%	539
b. Program staff actively and continuously consult and involve youth.	0.7%	1.7%	42.5%	53.7%	1.3%	536
c. Program staff provide structured and planned activities explicitly designed to help youth to get to know one another.	1.5%	1.9%	42.0%	52.6%	2.0%	538
d. Program staff provide opportunities for youth to lead activities.	1.3%	5.2%	49.2%	41.3%	3.0%	535
e. Program staff provide opportunities for youth to help or mentor other youth in completing a project or task.	1.5%	4.7%	47.5%	41.7%	4.7%	535
f. Program staff provide opportunities for the work, achievements, or accomplishments of youth to be publicly recognized.	1.5%	3.4%	45.0%	45.9%	4.3%	536

Staff were further asked to provide their perceptions regarding opportunities staff create for youth autonomy and choice. As Table 9 shows, the responses were somewhat lower on the agreement scale, though respondents still tended to agree with each individual practice listed. Staff expressed the highest levels of agreement in response to the item "staff are effective at finding ways to provide youth with meaningful choices when delivering activities," whereas agreement was less strong (though still strong overall) for "staff provide ongoing opportunities for youth to reflect on their experiences."

Table 9. Responses Concerning Provision of Opportunities for Youth Autonomy and Choice

Please rate the extent to which you agree or disagree with the following statements regarding all staff that work with students in this program:	Strongly Disagree	Disagree	Agree	Strongly Agree	Not Sure	N Responses
a. Program staff provide ongoing opportunities for youth to reflect on their experiences (e.g., formal journal writing, informal conversational feedback).	1.3%	7.6%	51.2%	34.0%	5.9%	539
b. Program staff are effective at finding ways to provide youth with meaningful choices when delivering activities.	1.3%	1.9%	47.5%	46.5%	2.8%	535
c. Program staff are effective at providing youth with opportunities to set goals and make plans within the confines of the program.	1.3%	5.2%	46.2%	43.9%	3.4%	535
d. Program staff ask for and listen to student opinions about the way things should work in the program.	1.9%	6.9%	45.3%	41.9%	4.1%	537

Respondents were next asked a set of questions concerning general program design. The respondents were instructed to indicate how *frequently* they led or participated in program activities that were exemplified by a particular characteristic. For each characteristic, fewer than half of the respondents selected "always," though, generally, around three-quarters of respondents answered either "always" or "frequently." As Table 10 shows, staff were most likely to select "rarely" or "sometimes" when asked about leading or participating in activities that were "structured to respond to youth feedback on what the content or format of the activity should be" or "informed by the expressed interests, preferences, and/or satisfaction of participating youth."

Table 10. Responses Concerning Program Design

How often do you lead or participate in program activities that are…	Rarely	Sometimes	Frequently	Always	N Responses
a. Based on written plans for the session, assignments, and projects?	4.1%	14.0%	41.9%	39.9%	534
b. Well planned in advance?	1.1%	10.3%	44.8%	43.8%	534
c. Tied to specific learning goals?	3.0%	12.5%	39.1%	45.3%	534
d. Meant to build upon skills cultivated in a prior activity or session?	2.8%	14.7%	45.1%	37.4%	532
e. Explicitly meant to promote skill building and mastery in relation to one or more state standard?	3.8%	14.4%	42.2%	39.6%	533
f. Explicitly meant to address a specific developmental domain (e.g., cognitive, social, emotional, civic, physical)?	2.5%	14.3%	44.2%	39.1%	530
g. Structured to respond to youth feedback on what the content or format of the activity should be?	2.8%	22.0%	46.0%	29.2%	531
h. Informed by the expressed interests, preferences, and/or satisfaction of participating youth?	1.9%	16.4%	47.5%	34.2%	535

Respondents were then asked about program linkages to the school day using an agreement scale (see Table 11). During this questioning, respondents were less likely to indicate strong agreement, with many respondents choosing “agree.” Areas of strongest agreement (about three-quarters of respondents or higher) were “I know whom to contact at my students’ day school if I have a question about their progress or status” and “The activities I provide in the afterschool program are tied to specific learning goals that are related to the school-day curriculum.” Areas of strongest disagreement (about a quarter of respondents disagreeing) were “I meet regularly with school-day staff not working in the afterschool program to review the academic progress of individual students” and “I participate in parent–teacher conferences to provide information about how individual students are faring in the afterschool program.” As many as a quarter of respondents per item reported that the given statement was “not relevant to my role in the program.”

Table 11. Responses Concerning Communication and Linkages to the School Day

Please rate the extent to which you agree or disagree with the following statements regarding linkages to the school day:	Strongly Disagree	Disagree	Agree	Strongly Agree	Not Relevant to My Role in the Program	Not Sure	N Responses
a. On a week-to-week basis, I know what academic content will be covered during the school day with the students I work with in the afterschool program.	4.4%	12.4%	34.0%	28.5%	18.8%	1.9%	526
b. I coordinate the content of the afterschool activities I provide with my students' school-day homework.	3.3%	17.2%	31.2%	20.8%	25.0%	2.5%	523
c. I know whom to contact at my students' day school if I have a question about their progress or status.	1.3%	4.4%	31.4%	51.4%	10.1%	1.3%	523
d. The activities I provide in the afterschool program are tied to specific learning goals that are related to the school-day curriculum.	1.9%	8.2%	37.6%	35.4%	14.8%	2.1%	526
e. I use student assessment data to provide different types of instruction to students attending my afterschool activities based on their ability level.	2.1%	16.6%	32.4%	23.7%	21.4%	3.8%	524
f. I help manage a formal 3-way communication system that links parents, program, and day-school information.	2.9%	10.5%	35.6%	24.3%	23.8%	3.0%	526

Please rate the extent to which you agree or disagree with the following statements regarding linkages to the school day:	Strongly Disagree	Disagree	Agree	Strongly Agree	Not Relevant to My Role in the Program	Not Sure	N Responses
g. I participate in regular, joint staff meetings for afterschool and regular school-day staff where steps to further establish linkages between the school day and afterschool are discussed.	2.7%	12.7%	32.6%	34.1%	16.4%	1.5%	519
h. I meet regularly with school-day staff not working in the afterschool program to review the academic progress of individual students.	3.8%	19.6%	32.2%	18.3%	24.0%	2.1%	525
i. I participate in parent–teacher conferences to provide information about how individual students are faring in the afterschool program. (Note: If you are a school-day teacher, please respond to this question in relation to students you do not have in your school-day classroom).	3.6%	20.3%	25.7%	20.2%	28.7%	1.5%	526

Staff were also asked about their use of data, with question responses including “do not receive,” “occasionally use,” and “often use.” Respondents were also given the option “not relevant to my role in the program.” Table 12 shows that a plurality of staff indicated they do not receive individual student academic plans or standardized test scores, whereas a majority of respondents reported often or occasionally using student grades and receiving input from students’ school-day teachers.

Table 12. Responses Concerning Data Use

Please indicate whether you receive each of the following, and to what extent you use it in planning for the activities you provide:	Do not Receive	Occasionally Use	Often Use	Not Relevant to My Role in the Program	N Responses
a. Individual student academic plans	34.7%	18.2%	22.4%	24.7%	527
b. Students' standardized test scores	37.1%	19.4%	15.0%	28.5%	526
c. Students' grades	26.2%	19.9%	29.6%	24.3%	523
d. Input from students' day-school teachers	16.6%	25.0%	40.0%	18.5%	525
f. Other	42.6%	7.9%	17.8%	31.7%	101

Regarding service delivery practices that enable youth to lead activities, work collaboratively, choose their projects, and so on, survey respondents were asked how available these types of opportunities were to students. As Table 13 shows, response options included "never available," "available occasionally in some classes or activities," "available regularly in most classes or activities," and "always available." Respondents were instructed to consider only those activities they led themselves. Most respondents indicated that each practice asked about was at least available occasionally. The practice least available was "make formal presentations to the larger group of students," for which 9.1% of respondents indicated that such opportunities were never available.

Table 13. Responses Concerning Service Delivery Practices

How often are students participating in the activities ***you*** provide in the program afforded the following types of opportunities:	Never Available	Available Occasionally in Some Classes or Activities	Available Regularly in Most Classes or Activities	Always Available	N Responses
a. Work collaboratively with other students in small groups	1.9%	10.2%	45.7%	42.2%	519
b. Have the freedom to choose what activities or projects they are going to work on or participate in	3.7%	28.8%	38.0%	29.5%	518
c. Work on group projects that take more than one day to complete	5.4%	23.3%	43.9%	27.4%	519
d. Lead group activities	5.2%	30.9%	38.4%	25.5%	518
e. Provide feedback on the activities they are participating in during time set aside explicitly for this purpose	4.7%	21.9%	43.9%	29.5%	515

How often are students participating in the activities ***you*** provide in the program afforded the following types of opportunities:	Never Available	Available Occasionally in Some Classes or Activities	Available Regularly in Most Classes or Activities	Always Available	N Responses
f. Participate in activities that are specifically designed to help students get to know one another	4.4%	25.5%	42.9%	27.2%	518
g. Make formal presentations to the larger group of students	9.1%	41.0%	30.9%	19.0%	517

Respondents were asked further questions about specific program practices relating to youth ownership, with agreement scale response options. Table 14 shows that most respondents indicated "agree" to the different practices, with "youth are afforded opportunities to take responsibility for their own program" and "youth have the opportunity to set goals for what they want to accomplish in the program" being endorsed with the highest levels of overall agreement (both "agree" and "strongly agree"). However, around a quarter of respondents disagreed ("disagree" or "strongly disagree") that "youth make choices about *how* content is covered in program offerings" or that "youth help create rules and guidelines for the program."

Table 14. Responses Concerning Youth Ownership

Please indicate your level of agreement with the following statements about how your students build ownership of the program:	Strongly Disagree	Disagree	Agree	Strongly Agree	Not Sure	N Responses
a. Youth are afforded opportunities to take responsibility for their own program.	1.7%	7.3%	52.0%	30.4%	8.6%	523
b. Youth have the opportunity to set goals for what they want to accomplish in the program.	1.3%	7.9%	53.6%	31.5%	5.8%	521
c. Youth help make plans for what activities are offered at the program.	1.9%	10.2%	53.0%	26.6%	8.3%	519
d. Youth make choices about *what* content is covered in program offerings.	2.1%	16.9%	49.6%	19.4%	11.9%	520
e. Youth make choices about *how* content is covered in program offerings.	1.9%	23.8%	47.1%	15.0%	12.1%	520
f. Youth help create rules and guidelines for the program.	3.9%	19.3%	45.0%	21.6%	10.2%	518

Respondents were then asked a series of questions having to do with internal program communication, with response options on a frequency scale concerning how often the item in question takes place ("never," "a couple times per year," "about once a month," and "nearly every week"). For this question, responses were more varied than in some of the other question areas, with respondents generally spread across the range of response options. As Table 15 shows, in terms of frequent communication practices, about 74% of program staff reported engaging with other staff at least monthly to "share ideas on how to make programming more engaging for participating students." Similarly, about 78% of staff reported engaging at least monthly with other program staff to "share experiences and follow up about individual youth." In terms of infrequent practices, however, a large percentage of respondents (38%) reported that they never "observe other afterschool staff delivering programming to provide feedback on their practice."

Table 15. Responses Concerning Internal Communication

How frequently do you engage in the following tasks with other staff working in the afterschool program:	Never	A Couple of Times per Year	About Once a Month	Nearly Every Week	N Responses
a. Conduct program planning based on a review of program data	17.2%	27.0%	30.6%	25.2%	523
b. Use evaluation data to set program improvement goals	20.0%	29.2%	32.1%	18.7%	520
c. Discuss progress on meeting program improvement goals	11.3%	26.0%	35.1%	27.6%	515
d. Observe other afterschool staff delivering programming in order to provide feedback on their practice	38.4%	19.0%	22.3%	20.2%	515
e. Conduct program planning in order to meet specific learning goals in coordinated ways across multiple activities	17.3%	24.0%	31.3%	27.3%	520
f. Share ideas on how to make programming more engaging for participating students	5.4%	20.6%	30.2%	43.8%	520
g. Share experiences and follow up about individual youth	6.0%	15.7%	28.8%	49.5%	517
h. Receive feedback from school-day teachers and/or administrators on how the program could better support student learning needs	18.7%	20.8%	29.8%	30.7%	514
i. Participate in training and professional development on how to better serve youth	8.9%	45.5%	30.6%	14.9%	516
j. Discuss current research-based instructional practices	18.8%	33.1%	31.2%	16.9%	510

Finally, respondents were asked a series of questions about communication with parents, with respondents provided options according to frequency ("never," "sometimes," and "frequently"). As Table 16 shows, most staff reported that they *frequently* "send materials about program offerings home to parents" (58%). The type of communication with the largest "never" response was "ask for input from parents on what and how activities should be provided."

Table 16. Responses Concerning Parent Communication

How often do you or other center staff:	Never	Sometimes	Frequently	N Responses
a. Send materials about program offerings home to parents	7.3%	34.4%	58.3%	523
b. Send information home about how the student is progressing in the program	16.4%	51.8%	31.7%	517
c. Hold events or meetings to which parents are invited	11.6%	40.5%	48.0%	519
d. Have conversations with parents over the phone	14.2%	46.6%	39.3%	522
e. Meet with one or more parents	11.7%	51.2%	37.0%	521
f. Ask for input from parents on what and how activities should be provided	25.3%	50.5%	24.2%	521
g. Encourage parents to participate in center-provided programming meant to support their acquisition of knowledge or skills	17.8%	40.8%	41.4%	517
h. Encourage parents to participate in center-provided programming with their children	14.9%	39.9%	45.2%	516

Summary of Staff Survey Data

In summary, respondents tended to be well educated (with nearly all having at least some college, and approximately four of every five staff having at least a college degree) and tended to respond positively on most if not all scales. Respondents strongly indicated that provision of enrichment opportunities was a primary goal (with nearly 90% selecting this as a primary goal), and staff generally reported high levels of staff efficacy (e.g., staff actively consulting youth, providing structured and planned activities).

Despite the many positive responses, staff responses also indicated areas that could be improved. Respondents indicated less *agreement* about whether staff provide youth with ongoing opportunities for reflection, for example, and were at least somewhat more likely to respond "rarely" or "sometimes" concerning participating in or leading activities structured to respond to youth feedback. In this same vein, approximately a quarter of respondents indicated that youth do not have opportunities to determine how content is covered in program offerings or opportunities to help create rules and guidelines for the program. Of course, staff did report tying activities to specific learning goals, and staff generally indicated that activities that enable collaborative work or types of youth leadership are provided at least occasionally.

In addition, roughly a quarter of respondents indicated they do *not* meet regularly with school-day staff to discuss academic progress, and questions concerning communication about the school day yielded many answers in the “agree” column (rather than “strongly agree”). Frequent (“often”) use of data such as academic plans, test scores, grades, and teacher input was reported by only a minority of respondents. Furthermore, though internal program communication was reported as generally strong (with approximately three quarters of respondents indicating modestly frequent communication with peers or their team, depending on communication type), a large minority (38%) indicated that they never observe other staff delivering activities (for feedback purposes). Also, a full quarter of staff indicated their program “never” asks for parent feedback. Noting again that these responses were provided by non-new grants, these findings indicate some possible areas for profitable professional development.

Chapter 5. Assessing 21st CCLC Program Outcomes

Another primary objective of the statewide evaluation was to understand the relationship between participation in 21st CCLC–funded programs and student outcomes. Employing program participation and outcome data associated with the 2015–16 programming period, a series of analyses were undertaken to assess the extent of program impact on state assessment results, truancy rates, and retention (or grade promotion). These analyses were based on a rigorous quasi-experimental design that compared academic outcomes of 21st CCLC program participants with matched nonparticipating students using a propensity score matching (PSM) approach. In addition to comparing participants and nonparticipants, the analyses also compared the outcomes for high-attending students to those of students who attended but did so infrequently. This section describes the method and results.

Method of Analysis

The evaluation team employed a quasi-experimental research design to examine the effect of participating in 21st CCLC programming on students' reading and mathematics achievement measured by New Jersey reading and mathematics assessments as well as the number of truant days and the grade retention status. The goal of this analysis was to assess the extent to which there is evidence that students participating in services and activities funded by 21st CCLC demonstrate better outcomes compared with students not participating in the program, specifically with respect to

- Higher academic achievement in reading/language arts and mathematics
- Lower truancy and retention rates

The study compared the outcomes of students who participated in 21st CCLC programs to similar students who did not participate using a propensity score stratification approach. Comparisons based on participation (versus nonparticipation) and high attendance (versus low attendance) were defined in a number of ways. Participation was analyzed with the following treatment and comparison conditions:

- Truancy and retention for students attending at least 30 days in comparison with nonattenders;
- Truancy and retention for students attending at least 60 days in comparison with nonattenders;
- Mathematics achievement scores for students attending at least 30 days and logging at least 30 hours of mathematics-focused programming;
- Mathematics achievement scores for students attending at least 60 days and logging at least 30 hours of mathematics-focused programming;
- Reading achievement scores for students attending at least 30 days and logging at least 30 hours of reading-focused programming; and,
- Reading achievement scores for students attending at least 60 days and logging at least 30 hours of reading-focused programming.

In addition to these analyses that focused on participants in comparison with nonparticipants, comparisons were made between high-attending and low-attending students (within the population of students who attended 21st CCLC programming at least one day during 2015–16). The impact of high attendance was analyzed using the following treatment and comparison conditions:

- Truancy and retention for students attending at least 60 days in comparison with students attending between one and 29 days;
- Mathematics achievement scores for students attending at least 60 days and logging at least 60 hours of mathematics-focused programming in comparison with students attending between one and 29 days and logging 30 or fewer hours of mathematics-focused programming; and,
- Reading achievement scores for students attending at least 60 days and logging at least 60 hours of reading-focused programming in comparison with students attending between one and 29 days and logging 30 or fewer hours of reading-focused programming.

These definitions of "treatment" were determined to ensure that the comparison of program effects was based on students who received a significant dose of 21st CCLC programming.

Accounting for Selection Bias

In any evaluation of a program in which participants are not randomly assigned to participate, the problem of selection is paramount. It is likely that students who participate in 21st CCLC programming are different from those who do not attend. Differences can bias estimates of program effectiveness because they make it difficult to disentangle preexisting differences between students who attended the program and those who did not from the effect of attending the program. In general, the research found that students who attended the program tended to be higher achieving students than those who did not prior to the start of the current academic year. The quasi-experimental approach outlined here, PSM, is a method for mitigating that existing bias in program effect (i.e., if one were to simply compare the students who attended and those who did not).

PSM is a two-stage process designed to address this problem. In the first stage, the probability that each student would participate in the 21st CCLC program was modeled on available observable characteristics. By modeling selection into the program, this approach allowed comparison of participating and nonparticipating students who would have had a similar propensity to select the program based on observable characteristics that were available in the data received from New Jersey. In the second stage, the predicted probability of participation was used to model student outcomes while accounting for selection bias. Using balanced pretreatment group differences in observed covariates allowed a propensity score stratification and marginal mean weighting approach (Hong & Hong, 2009).

Stage 1: Creation of the Control Group. The outcome of interest in modeling propensity scores is treatment status (1 for students in treatment, 0 for nontreatment students). To account for this binary outcome, logistic regression was used to model the logit (or log-odds) of student group assignment status. Examples of student-level variables used to fit the propensity score models included the following:

- Prior achievement in reading and mathematics;

- Prior truancy and retention; and,
- Student demographic information, including gender, race/ethnicity, socioeconomic status, special education status, and limited English proficiency status.

Appendix B contains a full list of the student-level variables used in the models.

In addition to the student-level variables, the propensity score model also included school variables (also noted in the Appendix) that added information about the school a student attended (to account for school-based contextual differences, which may account for differences in the propensity for a student to participate). Data were not available for each of these covariates for all students. To account for this, indicator variables were used to model the relationship between the pattern of missing data and propensity to participate in the program (Rosenbaum & Rubin, 1984). The propensity score model was fit separately for each grade (Grades 4–8) and separately for each definition of treatment (six treatment definitions based on participation and three treatment definitions based on attendance/programming). The final propensity score models for each grade were checked to ensure that the analysis sample was balanced across relevant covariates. The propensity score models all produced comparison samples, which were balanced with the treatment across all student and school covariates. This result indicates that the treatment and comparison groups had no significant differences from one another (prior to treatment).

Results

As outlined in the previous section, the analyses focused on nine different definitions of treatment/comparison contrasts. Six of these contrasts were focused on participants and nonparticipants, and three were focused on high and low attenders (within the population of participants). The results are presented first for participants versus nonparticipants and then by the results for high versus low attendance.

Impact Based on 21st CCLC Program Participation

The first set of analyses examined the impact of participation in 21st CCLC programming on truancy, retention, and achievement in mathematics and reading. These analyses focused on two levels of participation: 30 or more days and 60 or more days. For the mathematics and reading achievement analyses, students were considered to be in the treatment group only if they also had received at least 30 hours of subject-specific programming. The comparison groups for all analyses are a matched group of students who did not participate at all in 21st CCLC programming in 2015–16.

Truancy and Retention. Attending 60 or more days had a statistically significant, negative effect on truancy (reduced truancy) and a statistically significant negative effect on retention (reduced retention). In terms of truancy, youth in the 60-or-more-days group had a rate of truancy that was .976 times that of nonparticipants (i.e., a very small difference), whereas the odds ratio for retention (comparing youth attending 60 or more days with nonparticipants) was .876 (about what was observed in the impact analysis of 2013–14 data, the last impact analysis conducted). There were no statistically significant differences for 30 or more days, indicating that effects of 21st CCLC programming on truancy and retention, at least at this level of analysis, are only observable at attendance thresholds higher than 30 days.[7] See Table 17.

What Are Odds Ratios?

Retention results are described in terms of "**odds ratios**." Odds, like in betting, can be expressed as *1:1* (even odds), *9:1 against* (not great odds), and so on. An odds *ratio* is consequently fairly simple. First, the odds for an individual to be retained a grade are calculated for both treatment and nontreatment groups. Then, these two sets of odds are expressed as a ratio.

Not at all clear? A hypothetical example: The odds of being retained might be 99:1 for nontreatment (one in 100 is retained), and 199:1 for treatment (one in *200* is retained). In this case (again, hypothetical), the odds ratio for the treatment group would be about .5 because the odds of retention for the treatment group are roughly half what they are for the nontreatment group.

Note that a "**rate ratio**" is something different and is used to describe the difference in truancy levels. Rate ratios are easier: a rate ratio of, say, *.94*, means the truancy rate for the treatment group was about 94% that of the nontreatment group (i.e., better).

Table 17. Pooled Impact of Participation on Truancy and Grade Retention

Outcome	Treatment	Effect Size	Standard Error	p	Students	
					Treatment	Comparison
Truancy	30+ days	–0.002	0.004	0.602	7,063	115,428
	60+ days	–0.024	0.003	<.001***	5,406	111,595
Retention	30+ days	0.000	0.000	0.111	7,063	115,428
	60+ days	–0.132	0.001	<.001***	4,065	94,534

Note. *SE* = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Running the same analysis at the individual grade level, the reduction in truancy seemed to be driven by negative (i.e., the desired direction), statistically significant effects of participation in Grades 6 through 8. Attending 30 or more and 60 or more days led to significantly lower truancy rates in both Grades 6 and 8, whereas in Grade 7 attending 60 or more days led to significantly lower truancy rates. The effect sizes for the 60-or-more-days group in particular were larger than those noted above, with rate ratios of .848, .896, and .930 for Grades 6, 7, and 8, respectively. As Table 18 shows, these rate ratios are not quite as high as were observed for 2013–14.

[7] It is important to note that retention rates were very low in general, so there is little variance with which to determine a relationship between participation and retention. Also, the impact analysis in 2013–14 used 70 or more days as a comparison group rather than 60 days, so the analysis differs in that respect.

Table 18. Grade-Level Impact of Participation on Truancy

Grade	Treatment	Effect Size	Standard Error	p	Students	
					Treatment	Comparison
4	30+ days	0.020	0.016	0.211	1,626	17,265
	60+ days	0.017	0.018	0.333	1,341	17,061
5	30+ days	0.027	0.014	0.061	1,765	19,655
	60+ days	–0.027	0.017	0.098	1,435	19,014
6	30+ days	–0.042	0.015	0.006**	1,527	26,117
	60+ days	-0.152	0.019	<0.001***	1,101	26,147
7	30+ days	0.002	0.017	0.882	1,224	27,038
	60+ days	–0.104	0.020	<0.001***	910	26,808
8	30+ days	–0.037	0.019	0.045*	921	25,353
	60+ days	–0.070	0.024	0.004**	619	22,565

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

As Table 19 shows, similar to truancy, the significant overall reduction in grade retention for participants seemed to come from students in Grades 6 through 8. Significant reductions in the odds of retention were seen for Grade 6 students attending 30 or more or 60 or more days (odds ratios of .441 and .270, respectively). In Grade 7, there was a significant difference based on attending 60 or more days (odds ratio of .876). In Grade 8, the difference was for students who attended at least 30 days (odds ratio .240).

Table 19. Grade-Level Impact of Participation on Grade Retention

Grade	Treatment	Effect Size	Standard Error	p	Students	
					Treatment	Comparison
4	30+ days	0.044	0.525	0.934	1,626	17,265
	60+ days	NA	NA	NA	1,341	17,061
5	30+ days	0.107	0.310	0.730	1,765	19,655
	60+ days	0.018	0.331	0.958	1,435	19,014
6	30+ days	–0.818	0.350	0.020*	1,527	26,117
	60+ days	–1.308	0.516	0.011*	1,101	26,147
7	30+ days	0.021	0.332	0.949	1,224	27,038
	60+ days	–0.132	0.001	<0.001***	910	26,808
8	30+ days	–1.429	0.588	0.015*	921	25,353
	60+ days	–1.152	0.634	0.069	619	22,565

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Mathematics and Reading Achievement. Participation had a small, but positive, statistically significant effect on mathematics achievement (with effect sizes of around .02 standard deviation units). The increase was statistically significant at both the 30-or-more-days and 60-or-more-days levels. To help place these results in context, Hill, Bloom, Black, and Lipsey (2008) found that, on average, the effect of a whole year of learning on assessment results (counting time both in and out of school) averaged 0.31 standard deviation units for reading and 0.42 standard deviation units for mathematics (see also Naftzger, Devaney, & Newman, 2015). Given the amount of time youth spent in 21st CCLC compared with time spent during the school day, the results shown in Table 20, though not large, are still significant. There was no statistically significant impact of participation on reading achievement, however, which is similar to findings in previous years.

What Are Standard Deviation Units?

All assessment test results are shown with **effect sizes** included in the tables. The effect size, *in these cases*, is expressed in **standard deviation (SD)** units. An SD is simply a measure of how widely dispersed data are: Low SD values indicate scores tightly grouped together, high SD values indicate scores widely distributed. On a typical bell curve, a single SD up *and* down from the mean score will cover about two thirds of all scores reported.

Table 20. Pooled Impact of Participation on Mathematics and Reading Achievement

Outcome	Treatment	Effect Size	Standard Error	p	Students	
					Treatment	Comparison
Mathematics	30+ days, 30+ hours mathematics	0.019	0.007	0.008**	4,517	93,674
	60+ days, 30+ hours mathematics	0.021	0.007	<0.001***	3,892	92,657
Reading	30+ days, 30+ hours reading	–0.002	0.006	0.745	4,949	99,934
	60+ days, 30+ hours reading	0.004	0.005	0.427	4,203	97,640

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Table 21 shows that the increase in mathematics achievement was driven by improvements at the elementary grades, with participation (at both the 30- and 60-day levels) having a statistically significant, positive impact on mathematics achievement for both Grades 4 and 5 (with effect sizes between about .07 and .09 standard deviation units). There was a statistically significant but *negative* impact on mathematics scores in Grade 8 at the 30-or-more-days level, however (of about –.116 standard deviation units). The overall number of treatment youth for Grade 8 was relatively small.

Table 21. Grade-Level Impact of Participation on Mathematics Achievement

Grade	Treatment	Effect Size	Standard Error	p	Students	
					Treatment	Comparison
4	30+ days, 30+ hours mathematics	0.091	0.029	0.002**	1,171	15,190
	60+ days, 30+ hours mathematics	0.092	0.031	0.003**	1,069	15,380
5	30+ days, 30+ hours mathematics	0.071	0.028	0.012*	1,248	17,468
	60+ days, 30+ hours mathematics	0.083	0.030	0.006**	1,118	16,905
6	30+ days, 30+ hours mathematics	–0.015	0.030	0.617	1,040	21,302
	60+ days, 30+ hours mathematics	0.008	0.034	0.819	839	20,559
7	30+ days, 30+ hours mathematics	0.058	0.035	0.098	715	23,724
	60+ days, 30+ hours mathematics	0.051	0.039	0.193	599	24,039
8	30+ days, 30+ hours mathematics	–0.116	0.052	0.025*	343	15,990
	60+ days, 30+ hours mathematics	–0.088	0.057	0.125	267	15,774

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Analyses of reading achievement were also run at the grade level (see Table 22). In this case, however, there were no significant relationships between participation and reading achievement at any grade level.

Table 22. Grade-Level Impact of Participation on Reading Achievement

Grade	Treatment	Effect Size	Standard Error	p	Students	
					Treatment	Comparison
4	30+ days, 30+ hours reading	0.025	0.023	0.288	1,192	14,890
	60+ days, 30+ hours reading	0.015	0.024	0.515	1,087	15,008
5	30+ days, 30+ hours reading	0.025	0.023	0.289	1,235	16,701
	60+ days, 30+ hours reading	0.030	0.024	0.220	1,105	16,162
6	30+ days, 30+ hours reading	–0.042	0.025	0.098	1,113	23,240
	60+ days, 30+ hours reading	0.005	0.027	0.857	893	23,235
7	30+ days, 30+ hours reading	–0.048	0.030	0.113	811	23,538
	60+ days, 30+ hours reading	–0.012	0.033	0.705	677	23,184
8	30+ days, 30+ hours reading	–0.001	0.034	0.987	568	21,565
	60+ days, 30+ hours reading	–0.013	0.038	0.726	441	20,051

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Impact Based on 21st CCLC Program Attendance (High Versus Low Attendance)

The second set of analyses examined the impact of high attendance on truancy, retention, and achievement in mathematics and reading. The data for these analyses were limited to those students who had attended a 21st CCLC program at least once during 2015–16. High attendance generally meant 60 or more days of attendance, and low attendance meant between one and 29 days of attendance. For the mathematics and reading achievement analyses, students were considered to be in the treatment group only if they also had received at least 60 hours of subject-specific programming.

The reason for this approach was to control, as best as possible, for any selection bias not otherwise caught by the participant-to-nonparticipant comparisons. That is, by comparing high-attending youth with low-attending youth, the analyses may have intrinsically controlled for variables *not* otherwise included in the propensity score matching process because both the group of interest and the comparison group are 21st CCLC attenders.

Truancy and Retention. In terms of truancy and retention, however, the analysis found no statistically significant relationship between higher 21st CCLC attendance and truancy rates or odds of grade retention. Likewise, conducting the same analysis at the individual grade level yielded no statistically significant findings. Although this certainly could mean that increased participation does not yield greater returns in terms of these outcomes, it also could mean that the split between the one-to-29-days group on the one hand and the 60-or-more-days group on the other is sufficient to reveal significant differences. It may be the case that significant outcome differences between 21st CCLC attendees at different attendance levels are more observable once youth reach 90 or more days, for example. However, at this point this idea is merely conjecture and is not readily testable; a difficulty in running these types of analyses is that there are simply fewer and fewer youth records available as the attendance bands are drawn further and further apart, translating directly into a loss of statistical power. See Tables 23, 24, and 25.

Table 23. Pooled Impact of High Attendance on Truancy and Grade Retention (60+ days versus 1 to 29 days)

Outcome	Effect Size	Standard Error	p	Students	
				Treatment	Comparison
Truancy	–0.002	0.002	0.228	5,325	2,287
Retention	0.000	0.000	0.896	3,099	1,414

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Table 24. Grade-Level Impact of High Attendance on Truancy (60+ days versus 1 to 29 days)

Grade	Effect Size	Standard Error	p	Students	
				Treatment	Comparison
4	0.020	0.045	0.655	1,330	265
5	–0.035	0.035	0.325	1,416	400
6	–0.039	0.037	0.300	1,064	475
7	0.029	0.036	0.411	896	608
8	–0.078	0.040	0.053	619	539

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Table 25. Grade-Level Impact of High Attendance on Grade Retention (60+ days versus 1 to 29 days)

Grade	Effect Size	Standard Error	p	Students	
				Treatment	Comparison
4	NA	NA	NA	1,330	265
5	0.034	0.816	0.967	1,416	400
6	4.741	8.019	0.554	1,064	475
7	NA	NA	NA	896	608
8	6.160	13.040	0.637	619	539

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Mathematics and Reading Achievement. Table 26 shows that there was a small but statistically significant positive relationship between high attendance and mathematics achievement (though the effect size was very small, at .006 standard deviation units). No significant relationship was found for reading achievement in this case.

Table 26. Pooled Impact of High Attendance on Mathematics and Reading Achievement (60+ days attendance and 60+ hours subject-specific programming versus 1 to 29 days attendance and 30 or fewer hours subject-specific programming)

Outcome	Effect Size	Standard Error	p	Students	
				Treatment	Comparison
Mathematics	0.006	0.003	0.045*	3,032	1,864
Reading	0.001	0.003	0.570	3,295	1,860

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Looking at achievement scores by grade level, high attendance was statistically significantly related to higher mathematics achievement in Grade 4, with a sizeable effect of .213 standard deviation units. In Grade 8, however, the opposite was true, with high attenders having statistically significant lower mathematics achievement (with an equally sizeable effect in the other direction, of –.212). However, the overall *n* counts for both treatment and comparison groups became quite small at the individual grade level, so these results should be treated with some caution. Again, there were no statistically significant effects observed for reading assessment results. See Tables 27 and 28.

Table 27. Grade-Level Impact of High Attendance on Mathematics Achievement (60+ days attendance and 60+ hours mathematics programming versus 1 to 29 days attendance and 30 or fewer hours mathematics programming)

Grade	Effect Size	Standard Error	p	Students	
				Treatment	Comparison
4	0.213	0.077	0.006**	840	226
5	0.070	0.067	0.302	898	354
6	0.120	0.068	0.075	684	383
7	0.044	0.076	0.559	439	493
8	–0.212	0.092	0.021*	171	408

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Table 28. Grade-Level Impact of High Attendance on Reading Achievement (60+ days attendance and 60+ hours reading programming versus 1 to 29 days attendance and 30 or fewer hours reading programming)

Grade	Effect Size	Standard Error	p	Students	
				Treatment	Comparison
4	0.088	0.059	0.136	873	230
5	–0.002	0.054	0.974	866	357
6	0.076	0.053	0.148	716	419
7	–0.028	0.064	0.665	532	468
8	–0.124	0.072	0.085	308	386

Note. SE = standard error; ***significant at 0.001; **significant at 0.01; *significant at 0.05.

Relationship Between Attendance and Outcomes for Students With an IEP

The third set of analyses focused on participating students (those who had attended 21st CCLC programming at least one day) who also had an IEP (see Table 29). A series of hierarchical linear regression models helped analyze the relationship between attendance and outcomes for these students. These models included all of the covariates used in the propensity score analyses implemented in the previous sections (i.e., prior achievement, truancy and retention, student demographics, and school demographics). To estimate the relationship between attendance and truancy and retention, these analyses included the number of days that students with an IEP attended 21st CCLC programming. The mathematics and reading outcomes were predicted by the number of hours students logged for subject-specific programming (i.e., mathematics or reading, depending on the outcome).

Almost no students in this group were retained in their current grade in 2015–16, so no retention model could be fit to the data. There were no other significant relationships between attendance and outcomes for this group of participating students.

Table 29. Relationship Between 21st CCLC Program Attendance and Student Outcomes for Participating Students With an IEP

Outcome	Effect Size	Standard Error	p	Students
Truancy	0.000	0.001	0.986	737
Retention	NA	NA	NA	NA
Mathematics	0.011	0.011	0.327	594
Reading	0.006	0.005	0.254	625

Summary of Results

The participant-versus-nonparticipant propensity score stratification approach employed seeks to minimize the impact of selection bias on the estimates of program impact. However, it is an untestable assumption that such models can fully account for selection bias, so a second series of analyses was conducted, this time between high-attending participants and low-attending participants, with the idea that selection bias would be more thoroughly controlled because all students in the model attended at least to some degree. However, even this approach is not equivalent to true random assignment, and, to the extent that other variables exist (not available

for this analysis) that predict student participation (or high levels of participation) in 21st CCLC and also are related to student achievement, these analyses may be limited. To that end, these analyses provide initial evidence about the impact of 21st CCLC on academic achievement, but, just as with past impact reports provided to NJDOE, they should not necessarily be considered equivalent to experimental studies, which have strong internal validity.

The results indicate that there is a positive impact of 21st CCLC programming on truancy rates for the 60-or-more-days group, though the effect size is very small. The impact of programming on truancy rates is larger for Grades 6 through 8, with rate ratios of .848, .896, and .930, respectively, for youth attending above the 60-day threshold. The effect of programming on truancy rates for Grades 6 and 8 were also statistically significant at the 30-day threshold, though small.

The overall impact of 21st CCLC programming on retention was apparent at the 60-or-more-days level, with an effect-size-odds ratio of .876. Looking at retention by grade levels, significant reductions were seen for Grade 6 students attending 30 or more or 60 or more days (odds ratios of .441 and .270, respectively), whereas for Grade 7 there was a significant difference based on attending 60 or more days (odds ratio of .876). In Grade 8, the difference was for students who attended at least 30 days (odds ratio of .240). However, the absolute incidence rate of retention is, on the whole, relatively low.

In terms of assessment score outcomes, there was evidence of a program effect on mathematics at the overall level, with about a .02 standard deviation unit difference between treatment and nontreatment groups at both the 30- and 60-or-more-days levels. This overall effect seemed to have been driven by effects witnessed for Grades 4 and 5 in particular, however, where the effect sizes were comparatively large (effect sizes between .07 and .09 standard deviation units). (As previously noted, Hill, Bloom, Black, and Lipsey [2008] found that, on average, the effect of a whole year of learning on assessment results [counting time both in and out of school] averaged 0.31 standard deviation units for reading and 0.42 standard deviation units for mathematics.) There was also a statistically significant *negative* effect observed for Grade 8 attendees at the 30-or-more-days level, however (though the overall *n* count for the treatment group was, in this case, relatively small, with a few hundred scores available). No effects on reading scores were detected, though this is in keeping with past impact analyses.

The analyses on high attenders versus low attenders did not yield many statistically significant findings, though there were statistically significant findings regarding mathematics for Grades 4 and 8 (positive in the case of the former, negative in the case of the latter). This analytic approach was adopted in an effort to better control otherwise unavailable variables that might predict 21st CCLC attendance. However, it may be the case that the attendance levels of the two groups (one to 29 days and 60 or more days) may not be sufficiently spread apart for statistical models to detect patterns of difference—at least not with confidence sufficient to rule out chance. This is presently an untestable hypothesis, however, because pushing the attendance brackets farther apart yields ever-smaller groups, which erodes analytic power.

Overall, the results are roughly in keeping with what has been observed in previous impact analyses conducted for NJDOE.

Chapter 6. Conclusions and Next Steps

Overall, the 21st CCLC program in New Jersey seems to be serving the population intended and is offering activities in keeping with New Jersey's 21st CCLC goals. In terms of staff surveys, respondent answers indicate that there may be opportunities for NJDOE to offer professional development guidance: Providing more opportunities for youth to determine content, enabling staff to observe each other (with the goal of helping each other learn), and increasing feedback requests from parents all are areas in which staff could potentially use support. Asking grantees about their interest in these areas might be a useful next step in that respect.

In terms of truancy, retention, and assessment outcomes, the program continues to have a modest if positive impact on truancy and retention, along with some impact observable on mathematics scores (though not for reading). The results do indicate a need for more sensitive outcome measures, however, which means the youth survey data collected for 2016–17 will be of high interest. Impact analyses conducted next year will incorporate the youth outcome measures.

References

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.* Washington, DC: Policy Studies Associates.

Black, A. R., Doolittle, F., Zhu, P., Unterman, R., & Grossman, J. B. (2008). *The evaluation of enhanced academic instruction in after-school programs: Findings after the first year of implementation* (NCEE 2008-4021). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/pdf/20084021.pdf

Durlak, J. A., Mahoney, J. L., Bohnert, A. M., & Parente, M. E. (2010). Developing and improving after-school programs to enhance youth's personal growth and adjustment: A special issue of AJCP. *American Journal of Community Psychology, 45*, 285–293.

Durlak, J. A., & Weissberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.

Granger, R. (2008). After-school programs and academics: Implications for policy, practice, and research. *Social Policy Report, 22*(2), 3–19. Retrieved from http://www.nccap.net/media/pages/Granger.pdf

Granger, R. C., Durlak, J. A., Yohalem, N., & Reisner, E. (2007). *Improving after-school program quality.* New York, NY: William T. Grant Foundation.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*, 172–177.

Hong, J., & Hong, Y. (2009). Reading instruction time and heterogeneous grouping in kindergarten: An application of marginal mean weighting through stratification. *Educational Evaluation and Policy Analysis, 31*(1), 54–81.

Hutchinson, C., Naftzger, N., & Miller, S. (2006). *South Carolina extended learning time study: Final report*. Chicago: Learning Point Associates.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. A., Snow, D., & Martin-Glenn, M. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*(2), 275–313.

Little, P. (2007). *The quality of school-age child care in after-school settings* (Research to Policy Connections, No. 7). New York, NY: National Center for Children in Poverty, Columbia University Mailman School of Public Health.

Naftzger, N., Bonney, C., Donahue, T., Hutchinson, C., Margolin, J., & Vinson, M. (2007). *21st Century Community Learning Centers (21st CCLC) analytic support for evaluation and program monitoring: An overview of the 21st CCLC performance data: 2005–06*. Naperville, IL: Learning Point Associates. Retrieved from http://www2.ed.gov/programs/21stcclc/2501yr3ppics.doc

Naftzger, N., Devaney, E., Newman, J. (2015). *National scan of 21st CCLC data, impact & quality improvement systems project: Findings report*. Chicago, IL: American Institutes for Research.

Naftzger, N., Nistler, M., Manzeske, D., Swanlund, A., Shields, J., Rapaport, A., Smith, C., Gersh, A., & Hallman, S. (2012). *Texas 21st Century Community Learning Centers: An interim evaluation report.* Naperville, IL: American Institutes for Research.

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.

State of New Jersey Department of the Treasury. (2013). Request for proposals: Statewide evaluation of the 21st Century Community Learning Centers program. Trenton, NJ: Author.

Vandell, D. L., Reisner, E. R., Brown, B. B., Dadisman, K., Pierce, K. M., Lee, D., & Pechman, E. M. (2005). *The study of promising after-school programs: Examination of intermediate outcomes in Year 2*. Madison, WI: Wisconsin Center for Education Research. Retrieved from http://childcare.wceruw.org/pdf/pp/year2_executive_summary_and_brief_report.pdf

Wilson-Ahlstrom, A., Yohalem, N., DuBois, D., & Ji, P. (2011). *From Soft Skills to Hard Data: Measuring Youth Program Outcomes*. Washington, DC: Forum for Youth Investment.

Wilson-Ahlstrom, A., & Yohalem, N. (with Pittman, K.). (2007). *Building quality improvement systems: Lessons from three emerging efforts in the youth-serving sector*. Washington, DC: Forum for Youth Investment, Impact Strategies.

Yohalem, N., Wilson-Ahlstrom, A., Fischer, S., & Shinn, M. (2009). *Measuring youth program quality: A guide to assessment tools* (2nd ed.). Washington, DC: Forum for Youth Investment.

Zief, S. G., Lauver, S., & Maynard, R. A. (2006). *The impacts of after-school programs on student outcomes: A systematic review for the Campbell Collaboration.* Oslo, Norway: Campbell Collaboration. Retrieved from http://www.campbellcollaboration.org/lib/download/58/

Appendix A. New Jersey 21st CCLC Staff Survey

Please rate the extent to which you agree or disagree with the following statements regarding all staff that work with students in this program:	Strongly Disagree	Disagree	Agree	Strongly Agree	Not Sure
a. Program staff listen to youth more than talk at them.	○	○	○	○	○
b. Program staff actively and continuously consult and involve youth.	○	○	○	○	○
c. Program staff provide structured and planned activities explicitly designed to help youth to get to know one another.	○	○	○	○	○
d. Program staff provide opportunities for youth to lead activities.	○	○	○	○	○
e. Program staff provide opportunities for youth to help or mentor other youth in completing a project or task.	○	○	○	○	○
f. Program staff provide opportunities for the work, achievements, or accomplishments of youth to be publicly recognized.	○	○	○	○	○

Please rate the extent to which you agree or disagree with the following statements regarding all staff that work with students in this program:	Strongly Disagree	Disagree	Agree	Strongly Agree	Not Sure
a. Program staff provide ongoing opportunities for youth to reflect on their experiences (e.g., formal journal writing, informal conversational feedback).	○	○	○	○	○
b. Program staff are effective at finding ways to provide youth with meaningful choices when delivering activities.	○	○	○	○	○
c. Program staff are effective at providing youth with opportunities to set goals and make plans within the confines of the program.	○	○	○	○	○
e. Program staff ask for and listen to student opinions about the way things should work in the program.	○	○	○	○	○

How often do you lead or participate in program activities that are…	Rarely	Sometimes	Frequently	Always
a. Based on written plans for the session, assignments, and projects?	○	○	○	○
b. Well planned in advance?	○	○	○	○
c. Tied to specific learning goals?	○	○	○	○
d. Meant to build upon skills cultivated in a prior activity or session?	○	○	○	○
e. Explicitly meant to promote skill building and mastery in relation to one or more state standard?	○	○	○	○
f. Explicitly meant to address a specific developmental domain (e.g., cognitive, social, emotional, civic, physical)?	○	○	○	○
g. Structured to respond to youth feedback on what the content or format of the activity should be?	○	○	○	○
h. Informed by the expressed interests, preferences, and/or satisfaction of participating youth?	○	○	○	○

Please rate the extent to which you agree or disagree with the following statements regarding linkages to the school day:	Strongly Disagree	Disagree	Agree	Strongly Agree	Not Relevant to My Role in the Program	Not Sure
a. On a week-to-week basis, I know what academic content will be covered during the school day with the students I work with in the afterschool program.	○	○	○	○	○	○
b. I coordinate the content of the afterschool activities I provide with my students' school-day homework.	○	○	○	○	○	○
c. I know whom to contact at my students' day school if I have a question about their progress or status.	○	○	○	○	○	○
d. The activities I provide in the afterschool program are tied to specific learning goals that are related to the school-day curriculum.	○	○	○	○	○	○
e. I use student assessment data to provide different types of instruction to students attending my afterschool activities based on their ability level.	○	○	○	○	○	○
f. I help manage a formal 3-way communication system that links parents, program, and day-school information.	○	○	○	○	○	○
g. I participate in regular, joint staff meetings for afterschool and regular school-day staff where steps to further establish linkages between the school day and afterschool are discussed.	○	○	○	○	○	○
h. I meet regularly with school-day staff not working in the afterschool program to review the academic progress of individual students.	○	○	○	○	○	○
i. I participate in parent–teacher conferences to provide information about how individual students are faring in the afterschool program. (NOTE: If you are a school-day teacher, please respond to this question in relation to students you do not have in your school-day classroom).	○	○	○	○	○	○

Please indicate whether you receive each of the following and to what extent you use it in planning for the activities you provide:	**Do Not Receive**	**Occasionally Use**	**Often Use**	**Not Relevant to My Role in the Program**
a. Individual student academic plans.	○	○	○	○
b. Students' standardized test scores.	○	○	○	○
c. Students' grades.	○	○	○	○
d. Input from students' day school teachers.	○	○	○	○
f. Other. Specify _____	○	○	○	○

How often are students participating in the activities *you* provide in the program afforded the following types of opportunities:	**Never Available**	**Available Occasionally in Some Classes or Activities**	**Available Regularly in Most Classes or Activities**	**Always Available**
a. Work collaboratively with other students in small groups.	○	○	○	○
b. Have the freedom to choose what activities or projects they are going to work on or participate in.	○	○	○	○
c. Work on group projects that take more than one day to complete.	○	○	○	○
d. Lead group activities.	○	○	○	○
e. Provide feedback on the activities they are participating in during time set aside explicitly for this purpose.	○	○	○	○
f. Participate in activities that are specifically designed to help students get to know one another.	○	○	○	○
g. Make formal presentations to the larger group of students.	○	○	○	○

Please indicate your level of agreement with the following statements about how your students build ownership of the program:	**Strongly Disagree**	**Disagree**	**Agree**	**Strongly Agree**	**Not Sure**
a. Youth are afforded opportunities to take responsibility for their own program.	○	○	○	○	○
b. Youth have the opportunity to set goals for what they want to accomplish in the program.	○	○	○	○	○
c. Youth help make plans for what activities are offered at the program.	○	○	○	○	○
d. Youth make choices about *what* content is covered in program offerings.	○	○	○	○	○
e. Youth make choices about *how* content is covered in program offerings.	○	○	○	○	○
f. Youth help create rules and guidelines for the program.	○	○	○	○	○

How frequently do you engage in the following tasks with other staff working in the afterschool program:	Never	A Couple of Times per Year	About Once a Month	Nearly Every Week
a. Conduct program planning based on a review of program data.	○	○	○	○
b. Use evaluation data to set program improvement goals.	○	○	○	○
c. Discuss progress on meeting program improvement goals.	○	○	○	○
d. Observe other afterschool staff delivering programming in order to provide feedback on their practice.	○	○	○	○
e. Conduct program planning in order to meet specific learning goals in coordinated ways across multiple activities.	○	○	○	○
f. Share ideas on how to make programming more engaging for participating students.	○	○	○	○
k. Share experiences and follow up about individual youth.	○	○	○	○
l. Receive feedback from school-day teachers and/or administrators on how the program could better support student learning needs.	○	○	○	○
m. Participate in training and professional development on how to better serve youth.	○	○	○	○
n. Discuss current research-based instructional practices.	○	○	○	○

How often do you or other center staff:	Never	Sometimes	Frequently
a. Send materials about program offerings home to parents.	○	○	○
b. Send information home about how the student is progressing in the program.	○	○	○
c. Hold events or meetings to which parents are invited.	○	○	○
d. Have conversations with parents over the phone.	○	○	○
e. Meet with one or more parents.	○	○	○
f. Ask for input from parents on what and how activities should be provided.	○	○	○
g. Encourage parents to participate in center-provided programming meant to support their acquisition of knowledge or skills.	○	○	○
h. Encourage parents to participate in center-provided programming with their children.	○	○	○

Appendix B. Variables Used in Propensity Score Matching Models

Student Level

Free or reduced-price lunch status
Limited English proficiency status
Prior-year truancy (cumulative days toward truancy—i.e., prior absences)
Prior-year grade promotion status
Prior-year mathematics achievement score (scale score)
Prior-year English language arts/reading achievement score (scale score)
Male or female indicator
Ethnicity (White, Black, Asian, Pacific Islander, American Indian)
Immigrant indicator
Home language: English status
Home language: Spanish status

School Level

Total enrollment
Total teaching staff
Number of full-time teachers
Average years of teacher experience
Highest teacher level of education: bachelor's degree
Highest teacher level of education: master's degree
Highest teacher level of education: doctoral degree
Title I mathematics status
Title I reading/language arts status
Title I science status
Percentage White
Percentage Black
Percentage Asian
Percentage Pacific Islander
Percentage American Indian
Percentage free or reduced-price lunch status
Percentage limited English proficiency status

ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.

AMERICAN INSTITUTES FOR RESEARCH®

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

www.air.org

Making Research Relevant

LOCATIONS

Domestic

Washington, D.C.
Atlanta, GA
Austin, TX
Baltimore, MD
Cayce, SC
Chapel Hill, NC
Chicago, IL
Columbus, OH
Frederick, MD
Honolulu, HI
Indianapolis, IN
Metairie, LA
Naperville, IL
New York, NY
Rockville, MD
Sacramento, CA
San Mateo, CA
Waltham, MA

International

Egypt
Honduras
Ivory Coast
Kyrgyzstan
Liberia
Tajikistan
Zambia

3067_11/17