Texas 21st Century Community Learning Centers

2012–13 and 2013–14
Combined Evaluation Report

Prepared for: Texas Education Agency

AIR®
American Institutes for Research®

July 2015

Table of Contents

List of Tables

List of Figures

List of Acronyms

21st CCLC	21st Century Community Learning Centers
ACE	Afterschool Centers on Education
AIR	American Institutes for Research
APT-O	Assessment of Program Practices Tool–Observation
CSM	Critical Success Model
ELA	English language arts
HLM	Hierarchical linear modeling
OCES	Observation of Child Engagement Scale
OST	Out-of-school time
PEIMS	Public Education Information Management System
POS	Point-of-service
PQA	Program Quality Assessment
PSM	Propensity Score Matching
RFP	Request for Proposal
RQ	Research Question
SPP	STAAR Pilot Project
STAAR	State of Texas Assessments of Academic Readiness
TAKS	Texas Assessment of Knowledge and Skills
TEA	Texas Education Agency
TX21st	TX21st CCLC Student Tracking System

Texas 21st Century Community Learning Centers

2012–13 and 2013–14 Combined Evaluation Report

July 2015

Prepared for:

Texas Education Agency

AMERICAN INSTITUTES FOR RESEARCH®

1120 East Diehl Road, Suite 200
Naperville, IL 60563-1486
800-356-2735 • 630-649-6500
www.air.org

CREDITS

American Institutes for Research
1120 East Diehl Road, Suite 200
Naperville, IL 60563-1486
Phone: 630-649-6500
www.air.org

Gibson Consulting Group
1221 South MOPAC Expressway, Suite 355
Austin, TX 78746
Phone: 512-328-0884
www.gibsonconsult.com

David P. Weikart Center for Youth Program Quality
301 W. Michigan Ave., Suite 200
Ypsilanti, MI 48197
Phone: 734-961-6900
www.cypq.org

Contributing Authors

American Institutes for Research
Elizabeth Devaney
Neil Naftzger
Feng Liu
Janet Levings
Sarah Cristobal
Matt Vinson
Michael Hutson
Mary Nistler

Gibson Consulting Group
Joseph Shields
Shana Shaw

David P. Weikart Center for Youth Program Quality
Barbara Hillaker
Samantha Hallman

Submitted to

Texas Education Agency
1701 North Congress Avenue
Austin, Texas 78701-1494
Phone: 512-463-9734

Executive Summary

Background and Context

A large body of research has shown that afterschool programs can have a positive impact on the young people who attend them, particularly young people from low-income communities. In fact, studies have shown that, when young people attend high-quality programs on a regular basis, improvement occurs in a variety of academic outcomes, including attendance, discipline referrals, achievement tests, and grades (Auger, Pierce, & Vandell, 2013; Kauh, 2011; Miller, 2003; Naftzger, Manzeske, Nistler, & Swanlund, 2013). Likewise, these programs can have an effect on social and emotional outcomes as well, particularly when the programs are focused explicitly on developing those skills (Durlak & Weissberg, 2007).

The 21st Century Community Learning Centers (21st CCLC) program, authorized under Title IV, Part B, of the Elementary and Secondary Education Act, as amended by the No Child Left Behind Act of 2001, attempts to provide academic enrichment opportunities during nonschool hours for children, particularly students who attend high-poverty and/or low-performing schools.[1] The federal grants are awarded to state education agencies, which, in turn, make competitive awards to eligible grantees to support afterschool and summer learning programs.[2] In July 2002, the federal government awarded Texas Education Agency (TEA) $24.5 million to fund TEA's first cohort of 21st CCLC grantees for the 2003–04 school year. This and subsequent federal funding resulted in 297 grants being awarded in Texas throughout eight funding cycles.[3] All centers funded by the Texas 21st CCLC program, known in Texas as the *Afterschool Centers on Education* (ACE),[4] are expected to provide programs and services designed to support student performance in the following areas: academic performance, school attendance, school behavior, promotion rates, and graduation rates.

In 2012, TEA created a supplemental grant program as part of its 21st CCLC program initiative, with funding beginning in the 2012–13 school year and continuing through the 2013–14 school year. This new grant program, called the *STAAR [State of Texas Assessments of Academic Readiness] Pilot Project* (SPP), was designed to provide current 21st CCLC grantees with additional funding to provide academic intervention in core subjects and help students at risk for academic failure to meet or exceed academic

[1] For more information, review the authorizing legislation as part of the Elementary and Secondary Education Act (2001), Title IV, Part B see http://www2.ed.gov/policy/elsec/leg/esea02/pg55.html.
[2] Grantees include local education agencies, nonprofit organizations, for-profit organizations, institutions of higher education, and city or county government agencies.
[3] A cycle represents a cohort of grantees that receive funding for five years. Cycle 5, for example, represents the fifth such cohort to receive funding since TEA began funding for this grant.
[4] In Texas, the 21st CCLC program has its own unique brand that communicates the characteristics of the program and creates statewide awareness so that all Texas centers can identify themselves as part of a bigger picture. Although 21st CCLC is the federal funding source, the programs in Texas are referred to as *Afterschool Centers on Education*, or *Texas ACE*. The term *ACE* will be used throughout the report to refer to the programs in Texas unless reference is made to the federal funding source, in which case the term *21st CCLC* will be used.

standards. Emphasis was placed on helping students improve performance on the state's new assessment of student learning, the State of Texas Assessments of Academic Readiness (STAAR). Grantees were awarded competitive grants to establish SPP programming in selected centers where they already had established ACE-only programs.[5]

Consistent with the requirements for ACE programs, the grantee agencies were required to develop SPP programs that were aligned with the ACE program's Critical Success Model (CSM), specifically the following:

- Innovative instructional techniques based on research and best practices be implemented.
- Adult advocates support student involvement in school.
- Preassessment and postassessment data be used to identify student needs and provide targeted interventions.
- Professional development be provided to staff to increase their effectiveness.

In November 2012, 15 21st CCLC grantee agencies were awarded SPP supplemental grants through a competitive process and established SPP programs in 46 centers where they also were implementing ACE programs. The grants were continued into the 2013–14 school year, but two Cycle 5 grantees ended their ACE programs at the conclusion of the 2012–13 school year, so only 13 grantees were included in the 2013–14 evaluation.

Statewide Evaluation: Year 1 (2010–11) and Year 2 (2011–12)

Beginning in fall 2010, TEA contracted with American Institutes for Research (AIR) and its partners at Gibson Consulting Group and the David P. Weikart Center for Youth Program Quality to conduct a statewide evaluation of the Texas ACE program, geared toward two primary research objectives TEA established for the project:

- **Research Objective 1:** Identify and describe innovative strategies and approaches implemented by successful 21st CCLC programs.

[5] *Grantee* refers to the organization that serves as the fiscal agent on the 21st CCLC grant, and *the center* refers to the physical location where grant-funded services take place. The centers have defined hours of operation and a dedicated staff; they are required to have a position akin to a site coordinator. Each ACE grantee must have at least one center and may have as many as 20 centers. With regard to SPP centers, there were not dedicated SPP centers specifically, even though they were sometimes referred to that way. Rather, SPP centers were ACE centers that included targeted interventions as the part of the programming that students were recruited to attend. Moreover, there were also ACE-only students and traditional ACE programming at a given SPP center that was available to all students enrolled in the center. The key distinction is that only certain ACE centers received specific supplemental grant funding to include SPP programming, either alongside other traditional ACE programming or as stand-alone programming.

- **Research Objective 2:** Conduct a statewide assessment of 21st CCLC programs, operations, participation in the program, and student achievement outcomes.

The results of the first two years of the evaluation are presented in two reports (Naftzger, Manzeske, Nistler, & Swanlund, 2012; Naftzger, Manzeske, Nistler, & Swanlund, 2013). Key findings from those two years were as follows:

- There were three instructional approaches found to be associated with high levels of student engagement: clarity of purpose, intentional use of time, and an active and interactive activity leader.
- Organizational practices associated with high levels of quality include intentional program design, staff development and collaboration, methods to monitor improvement, linkages to the school day, and community connections.
- There was some evidence of a connection between high-quality programs and high levels of participation.
- Higher levels of attendance (60 or more days) in 21st CCLC–funded programs were associated with higher levels of state assessment scores in reading/English language arts and mathematics performance, reduced disciplinary incidents and school-day absences, and supported grade promotion. In addition, a high level of point-of-service (POS) quality was associated with greater impact on reduced disciplinary incidents and grade promotion.
- For high school students, participation in an ACE program increased the likelihood of being promoted to the next grade level by 97 percent. There were similar findings for elementary and middle school students, but the magnitude was much smaller.

Statewide Evaluation: Year 3 (2012–13) and Year 4 (2013–14)

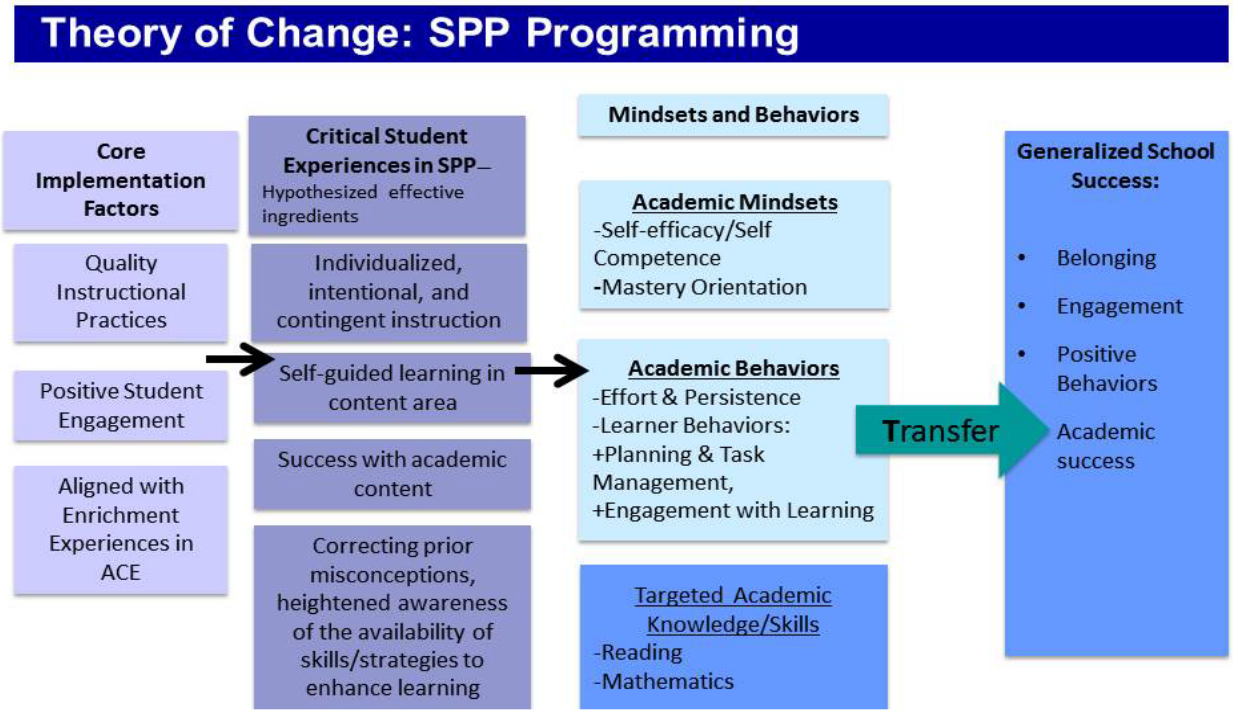
The introduction of the SPP program led to a change in the evaluation focus, although the scope remains the same. Beginning with the 2012–13 evaluation, AIR focused its activities and questions specifically on the SPP program as it compared with traditional ACE programming.

Linking Quality to Outcomes Through a Theory of Change

As noted previously, program and activity quality are expected to affect student engagement and, thus, student learning. Research supports the SPP theory of change that was developed by the evaluation team in partnership with TEA (Auger, Pierce, & Vandell, 2013; Durlak & Weissberg, 2007; Kataoka & Vandell, 2013; Kauh, 2011; Smith et al., 2012).

The SPP theory of change, depicted in Figure ES1, articulates the key facets of SPP implementation that contribute to the experiences youth may have in the program. A sequence of high-quality, engaging experiences across time will lead to students developing key beliefs and skills, both social-emotional and academic in nature. Improving these key mindsets and behaviors will, in turn, affect youth performance on key metrics during the school day.

Figure ES1. SPP Theory of Change



Research Questions

The evaluation of the SPP programs had six research questions (RQs) that continue to support the broader statewide evaluation objectives of impact and implementation and are designed to explore the validity of the SPP theory of change presented previously. Questions focus on describing the similarities and differences between ACE-only programs and SPP programs, particularly differences associated with program operations, activities, and participating students; examining the quality of SPP programming, variations in quality across programs, and the program operations that SPP grantees have implemented to support quality programming; and learning how and the extent to which the SPP programming impacts participating students. The six RQs that guided the 2012–13 and 2013–14 evaluation activities are as follows:

- **RQ 1:** How does SPP programming compare with ACE-only programming in centers administered by the same grantee?
- **RQ 2:** How do students participating in SPP programming differ from students who participate in ACE-only programming?
- **RQ 3:** How does the quality of delivery differ between SPP and ACE-only programming?
- **RQ 4:** What instructional and administrative practices lead to high student engagement?
- **RQ 5:** How engaged are young people in SPP and ACE-only programming? What is the relationship between quality of delivery and student engagement?
- **RQ 6:** What is the impact of SPP programming on students' academic mindsets and behaviors? How does this compare with ACE-only students?

A Summary of Preliminary Findings

The evaluation explores this theory of change by examining program quality, staff experiences in the program, youth mindsets and behaviors, and school-related outcomes. The preliminary findings across RQs are as follows. Some RQs were addressed only partially at this time because of the intention to conduct more rigorous impact analyses in 2014–15. The 2012–13 and 2013–14 evaluations were intended to be more descriptive and exploratory in nature in order to inform additional analyses in 2014–15.

RQ 1. How does SPP programming compare with ACE-only programming in centers administered by the same grantee?

Overall, findings from 2012–13 and 2013–14 indicate that **SPP programs were more academic in nature than were ACE-only programs**. This finding is what was expected given the purpose of the funding for SPP programming and the stated and explicit intention that SPP programs help students at risk for academic failure improve their skills. Specific findings that contributed to this overall conclusion include the following:

- *SPP programs hired more certified teachers.* In both years of study, a far higher percentage of SPP activity leaders (85 percent in 2012–13 and 81 percent in 2013–14) than ACE-only activity leaders (51 percent in 2012–13 and 42 percent in 2013–14) were credentialed teachers.
- *SPP students spent more time in academic activities. S*tudents participating in SPP activities spent much more of their time in academic activities—particularly in academic enrichment activities and tutoring (82 percent in 2012–13 and 66 percent in 2013–14 for SPP students compared with 45 percent in 2012–13 and 33 percent in 2013–14 for ACE-only students). Compared with ACE-only students, SPP students spent very little time in recreational activities (8 percent versus 33 percent in 2012–13 and 25 percent versus 33 percent in 2013–14). Students participating in SPP activities also spent much less of their time in homework help sessions compared to students participating in ACE-only activities. In 2013–14, SPP students began participating in recreation programs and other nonacademic enrichment to a larger degree, but there was still a substantial difference between the participation of the two types of students.
- *SPP activities used smaller groups and longer activities to support academic learning.* SPP activities typically met in smaller groups than did ACE-only activities, and SPP activity leaders served fewer students (on average six fewer) per week compared to ACE-only activity leaders. In addition, on average, SPP academic activities were 90 minutes long, compared with 60 minutes for ACE-only academic activities.
- *SPP and ACE-only activity leaders used different instructional strategies.* SPP activity leaders were more likely than ACE-only activity leaders to plan activities that explicitly promoted skill building related to state standards. They were also more likely than ACE activity leaders to use instructional strategies that explicitly addressed content knowledge—the use of computer-based learning programs, direct instruction, and the review and practice of concepts learned during the school day.

RQ 2. How do students participating in SPP programming differ from students who participate in ACE-only programming?

Analysis of student assessment data and youth survey data revealed that the students participating in SPP programming tended to be more academically at risk and less proficient in key academic mindsets and behaviors[6] than were their ACE-only peers at program onset. This finding is important because it indicates that the SPP program was successful in recruiting the types of students it intended to serve—that is, those at risk for academic failure. Key differences between the two types of students were as follows:

- *SPP students had higher academic needs than did ACE-only students.* Students who attended SPP programs in 2012–13 appeared to have higher academic

[6] Students were measured on the academic mindsets and behaviors through a survey given at program onset and again the spring after program completion. The survey measured competencies such as *Effort and Persistence*, *Learner Behaviors*, *Engagement*, and *Sense of Competence as a Learner*.

needs than did ACE-only students, as shown by their performance levels on the 2012 STAAR assessment. Specifically, between 19 percent and 36 percent fewer SPP students than ACE-only students met satisfactory proficiency levels on the various STAAR exams in 2012.

- *SPP students scored lower on the Mindsets and Behaviors Survey than did ACE-only students.* SPP students at all grade levels and across almost all constructs in both years scored lower than ACE-only students did on the Mindsets and Behaviors Survey at the program outset. In several cases, those differences were statistically significant. ACE-only students appeared, in particular, to be at a significantly higher level at the start of programming on constructs measured on the survey, including *Learner Behaviors* (at the K–3 and 4–12 grade levels) and *Academic Performance* (at the K–3 level).

RQ 3. How does the quality of delivery differ between SPP and ACE-only programming?

The evaluation team conducted analyses of the overall quality of both program types and found that **SPP programs and ACE-only programs had roughly the same level of overall quality**. The two types of programs were not the same, however, and the evaluation team found that **activity leaders differed in their administrative practices in several key ways**:

- *SPP activity leaders had more Internal Communication.* Activity leaders for SPP activities reported higher levels of *Internal Communication* than did those who led ACE-only program activities.
- *SPP activity leaders were more likely than ACE-only activity leaders to report* Program Challenges*.* The challenges they cited most frequently were inconsistent program attendance, student behavioral issues, and low levels of student interest in the activities. Challenges with technology also were noted.
- *SPP activity leaders focused on academic skills.* SPP activity leaders were more likely than ACE-only activity leaders to report implementing programs and activities focused on specific academic skills, particularly those connected to the state standards, and to direct students in *Academic Learning Activities* more frequently than were ACE-only activity leaders.

These differences, although not directly related to the point of service, affect the quality of program delivery. The last finding, in particular, is important because it connects to other findings in the report related to levels of youth engagement in explicitly academic programming.

RQ 4. What instructional and administrative practices lead to high student engagement?

By reviewing scores from observations of youth engagement, as well as youth responses to an engagement survey, and connecting those with observed and reported activity leader practices, the evaluation team was able to conclude that SPP **activities that used a *Learning Strategies* approach, a combination of computer-based and face-to-face**

delivery, and a low staff-to-youth ratio, were the most engaging to young people. The specific findings that supported this overall conclusion include the following:

- *Generalized* Learning Strategies *were associated with engagement.* SPP programs had two different approaches associated with curricular content. One was a *Learning Strategies* approach, which emphasized general learning strategies that were applicable across different content areas. A second common approach was a *Skills-Building* approach, which emphasized specific skills associated with a subject area. The *Learning Strategies* approach was associated with higher levels of quality than any other approach. That, combined with the findings in this report connecting quality and engagement, suggests that students may be more engaged when participating in activities using a *Learning Strategies* approach.
- *A combination of face-to-face and computer-based instruction was most engaging to students.* SPP activities also used two modes of delivery: computer-based delivery, in which students typically worked individually on an academic skills computer program, and face-to-face delivery, in which students worked mainly in small-group sessions facilitated by an activity leader. The majority of SPP programs used a combination of computer-based and face-to-face delivery modes. The observations of the SPP program activities, as well as the student engagement survey, indicated that engagement levels were higher among students in activities that combined computer-based and face-to-face delivery modes than among students in activities using only a single mode of delivery. Reported engagement was lowest among students in only computer-based activities.
- *Low staff-to-youth ratio was associated with youth engagement.* A low staff-to-youth ratio was associated with higher levels of youth engagement during observations of youth engagement.

RQ 5. How engaged are young people in SPP and ACE-only programming? What is the relationship between quality of delivery and student engagement?

An analysis of 2013–14 quality and engagement data showed that although overall there was not a significant difference between the quality of ACE-only and SPP programs, **youth-reported engagement was lower in SPP programs than in ACE-only programs**. The analysis of these two types of data also revealed that **overall, higher quality programs engaged students at higher levels, but in SPP programs, the relationship between program quality and youth-reported engagement was weaker, a finding that is hypothesized to be related to the provision of higher levels of academic content in SPP programs**. More specifically, the evaluation team found the following:

- *The Learning Strategies approach was associated with higher levels of quality than was any other approach*. The three SPP activities with the highest Program Quality Assessment (PQA) scores used a *Learning Strategies* approach for which there was an established curriculum.

- *SPP programs were consistently associated with lower levels of youth-reported engagement.* It is hypothesized that this finding is related to the provision of higher levels of academic content in these programs.
- *Youth development principles were associated with youth engagement.* SPP and ACE-only program activities that employed youth development principles like those outlined in the PQA (e.g., engaging youth in activities that lead to a tangible product, offering youth choice, and providing opportunities for reflection) had higher levels of youth self-reported engagement.

RQ 6. What is the impact of SPP programming on students' academic mindsets and behaviors? How does this compare with ACE-only students?

Findings related to this RQ were somewhat inconclusive, partly because the 2013–14 analysis was preliminary in nature at the time of this report. However, **there appears to be a small but positive impact of both SPP programming and potentially ACE-only programming on many of the measured academic mindsets and behaviors**. Again, this is important because improved mindsets and behaviors ultimately can lead to improved academic outcomes according to the SPP theory of change. Findings contributing to this conclusion include the following:

- *The 2012–13 results showed growth on the Academic Mindsets and Behaviors Survey.* The 2012–13 impact analysis showed some significant growth for students participating in SPP programming on the academic mindsets and behaviors measured on the student survey, including *Effort and Persistence* and *Learner Behaviors* for Grades K–3 and on *Effort and Persistence* for Grades 4–12. However, the amount of time between pre- and postsurveys was very short because SPP programming ran for only one semester in 2012–13, so students took the pretest in January and the posttest in April.
- *A* Learning Strategies *approach may support improvement in academic mindsets and behaviors.* Results suggest that students in Grades K–3 who participated in SPP programming using a *Learning Strategies* approach demonstrated a significantly[7] higher level of improvement on the Academic Mindsets and Behaviors Survey on *Effort and Persistence* and *Learning Behaviors* than did students in Grades K–3 who participated in SPP programming using a *Skills-Building* approach.
- *The 2013–14 results on the Academic Mindsets and Behaviors Survey were mixed.* In 2013–14, the analyses considered pre-post change in scores on the Academic Mindsets and Behaviors Survey for both types of students. Changes were positive and significant from fall to spring for both SPP and ACE-only students in Grades K–3 on all of the areas measured by the survey. The changes

[7] Statistically significant differences are defined in this report as those with a *p* value of 0.05 or below, which means that there is 5 percent probability (or less) of randomly observing a difference of this size or greater if no difference exists. A moderately significant difference would yield a *p* value less than 0.1, where there would be a 10 percent probability of observing a difference of this size by chance. Essentially, these two terms help to illustrate the degree of confidence the research team has in the findings.

from pre- to posttest were higher for ACE-only students, indicating that the program may have had a greater impact on ACE-only students' mindsets and behaviors than on SPP students'. For Grades 4-12, the analysis showed no improvement on the constructs for SPP students and a decrease in the mindsets and behaviors for ACE-only students—that is, their scores declined from fall to spring. These findings are inconclusive and should be viewed with extreme caution. The evaluation team has not conducted the more rigorous statistical analysis needed to answer some of the questions these preliminary findings raise. Right now they simply point to an interesting finding worth exploring in more detail in the 2014–15 analysis.

Taken together, these findings begin to tell a story about the validity of the SPP theory of change, a story that will be completed in 2014–15 during the final year of this evaluation. Interestingly, the findings from the Years 3 and 4 analyses indicate that the theory of change may be applicable to both SPP and ACE-only programs—that is, overall, the 2012–13 and 2013–14 analyses confirm that SPP and ACE-only programs are delivering activities that collectively are of average quality but that range from average to high quality and are roughly the same across both program types. This is an important first step in the theory of change. Students also report engagement when youth development practices are present, when the delivery is both face-to-face and computer based, and when activity leaders used a generalized learning approach rather than specific skill building. This finding was true across both types of programming. The findings also suggest a connection between quality and engagement, the first relationship posited in the theory of change. This finding was true for programs overall but was stronger in ACE-only programs. Finally, the evaluation findings show some evidence that both SPP and ACE-only programming had an impact on the mindsets and behaviors of some participants. So far these effects are small (2012–13) and inconclusive (2013–14) but show enough positive trends to warrant further study, particularly to understand better the progression from quality programming to engaged youth to outcomes related to mindsets, beliefs, and school performance.

Next Steps

The findings described are important for two reasons: (1) the 2012–13 findings were used to improve on methodology and inform 2013–14 data collection, and (2) the 2013–14 findings will be used to inform planned analyses for 2014–15. Each of these data uses is described in greater detail in the following sections.

Methodological Changes to the Evaluation

The 2012–13 evaluation procedures provided valuable experience, on which the evaluation team drew to improve 2013–14 data collection, including the following:

- The evaluation was conducted during a very short time in spring 2013. During 2013–14, evaluation activities spanned the full school year to allow more time for change to occur among students and to ease the burden on sites by having collection periods spread out.

- Analysis of the validity and reliability of the student survey instruments revealed that two scales on the Grade 4–12 student survey were not functioning well. The evaluation team removed those scales from the survey and replaced them with three new scales focused on a student's *Sense of Competence as a Learner.* In addition, the Grade K–3 survey had several items for which it was too easy for staff to rate the items highly. Those items were removed from the 2013–14 survey. Finally, a new scale was added to the Grade K–3 survey focused on academic performance and intended to be completed by school-day teachers.

Likewise, findings from the 2013–14 evaluation will inform 2014–15 evaluation plans. For example, the evaluation team will use preliminary findings from 2013–14 about quality and engagement to inform planned analyses related to the SPP theory of change. In particular, using 2013–14 data, the team will continue to explore how quality influences student engagement and what the mediating effects of engagement and academic mindsets and behaviors are on key outcomes. Much of the data and information presented in this combined, two-year report are preliminary and offer a first look at the SPP program and how it compares to the ACE-only programs in terms of program delivery, quality, students, and staff. In the coming year of the evaluation, the evaluation team will explore these findings in greater detail and conclude the exploration of the SPP theory of change begun with this report in order to provide recommendations for program implementation and professional development for staff involved in the programming.

Chapter 1: Introduction

A large body of research has shown that afterschool programs can have a positive impact on the young people who attend them, particularly young people from low-income communities. In fact, studies have shown that, when young people attend high-quality programs on a regular basis, improvement occurs in a variety of academic outcomes, including attendance, discipline referrals, achievement tests, and grades (Auger, Pierce, & Vandell, 2013; Kauh, 2011; Miller, 2003; Naftzger, Manzeske, Nistler, & Swanlund, 2013). Likewise, these programs can have an effect on social and emotional outcomes as well, particularly when the programs are focused explicitly on developing those skills (Durlak & Weissberg, 2007). Despite this research, across the United States, at least 8 million children are left alone and unsupervised each weekday afternoon (California Department of Education, 2002). Although there are a multitude of youth programs available, many of the programs are inaccessible to school-age children living in high-poverty communities. In Texas alone, there were just more than 5 million (5,151,925) students enrolled in public schools during the 2013–14 school year (Texas Education Agency [TEA], 2014). Of those students, two thirds are considered economically disadvantaged (TEA, 2014). Young people living in poverty are less likely to graduate than are their higher income peers as well as more likely to have lower school achievement levels (The Equity and Excellence Commission, 2013). Afterschool opportunities have been shown to correct this imbalance between lower and higher income youth (Auger, Pierce, & Vandell, 2013).

The 21st Century Community Learning Centers (21st CCLC) program, authorized under Title IV, Part B, of the Elementary and Secondary Education Act, as amended by the No Child Left Behind Act of 2001, attempts to correct this lack of access to critical afterschool opportunities for young people through the creation of community learning centers that provide academic enrichment opportunities during nonschool hours for children, particularly students who attend high-poverty and/or low-performing schools.[8] The federal grants are awarded to state education agencies, which, in turn, make competitive awards to eligible grantees to support afterschool and summer learning programs.[9] In July 2002, the federal government awarded TEA $24.5 million to fund TEA's first cohort of 21st CCLC grantees for the 2003–04 school year. This and subsequent federal funding resulted in 297 grants being awarded in Texas throughout eight funding cycles.[10] All centers funded by the Texas 21st CCLC program, known in Texas as the *Afterschool Centers on Education* (ACE),[11] are expected to provide

[8] For more information, review the authorizing legislation as part of the Elementary and Secondary Education Act (2001), Title IV, Part B at http://www2.ed.gov/policy/elsec/leg/esea02/pg55.html.

[9] Grantees include local education agencies, nonprofit organizations, for-profit organizations, institutions of higher education, and city or county government agencies.

[10] A cycle represents a cohort of grantees that receive funding for five years. Cycle 5, for example, represents the fifth such cohort to receive funding since TEA began funding for this grant.

[11] In Texas, the 21st CCLC program has its own unique brand that communicates the characteristics of the program and creates statewide awareness so that all Texas centers can identify themselves as part of a bigger picture. Although 21st CCLC is the federal funding source, the programs in Texas are referred to as *Afterschool Centers on Education*, or *Texas ACE*. The term *ACE* will be used throughout the report to refer to the programs in Texas unless reference is made to the federal funding source, in which case the term 21st CCLC will be used.

programs and services designed to support student performance in the following areas: academic performance, school attendance, school behavior, promotion rates, and graduation rates.

The ACE programs have been a long-established resource for providing Texas students with academic support and enrichment opportunities. In 2012, TEA identified the ACE programs as providers of even more focused academic support to Texas students who were at risk of academic failure, as measured by the state's new assessment of student learning, the State of Texas Assessments of Academic Readiness (STAAR). STAAR is a more rigorous assessment than the Texas Assessment of Knowledge and Skills (TAKS), which it replaced. STAAR was first administered to students in Grades 3–8, and the STAAR end-of-course assessment was administered first for Grade 9 in the 2011–12 school year. Each year thereafter, another high school grade was added (i.e., in 2012–13, Grade 10 was added). Although overall pass rates for the state were approximately 50 percent for mathematics and 67 percent for English language arts (ELA) in 2013, economically disadvantaged students passed at much lower rates than their more advantaged peers (by an average of 26 percentage points) (Center for Public Policy Priorities, 2013). This lower pass rate is notable, given the fact that ACE programs are designed to target students attending schools serving economically disadvantaged students and so have an opportunity to help those students perform better on state tests.

In 2012, TEA created a supplemental grant program as part of its 21st CCLC program initiative, with funding beginning in the 2012–13 school year and continuing through the 2013–14 school year. In the request for proposal (RFP) issued on October 3, 2012, the STAAR Pilot Project (SPP) was described as a "pilot initiative which will provide current 21st CCLC grantees with additional funding to increase their use of evidence-based interventions to support additional academic assistance in core subjects to help the students meet or exceed standards established by the new STAAR assessments." Grantees were awarded competitive grants to establish SPP programming in selected centers where they already had established ACE-only programs.[12]

Eligibility for SPP funding in 2012–13 was limited to grantees that had received 21st CCLC funds in Cycles 5, 6, and 7 and already had demonstrated success with respect to adherence to program requirements and the implementation of practices for

[12] Grantee refers to the organization that serves as the fiscal agent on the 21st CCLC grant, and the center refers to the physical location where grant-funded services take place. The centers have definedhours of operation and a dedicated staff; they are required to have a position akin to a site coordinator. Each ACE grantee must have at least one center and may have as many as 20 centers. With regard to SPP centers, there were not dedicated SPP centers specifically, even though programs were sometimes referred to that way. Rather, SPP centers were ACE centers that included targeted interventions as the part of the programming that students were recruited to attend. Moreover, there were also ACE-only students and traditional ACE programming at a given SPP center that was available to all students enrolled in the center. The key distinction is that only certain ACE centers received specific supplemental grant funding to include SPP programming, either alongside other traditional ACE programming or as stand-alone programming.

struggling students. Eligibility also was limited to grantees that had met certain program implementation criteria (i.e., requiring low to medium technical assistance, having no noncompliance incidents related to grant and program operation requirements, having met student participation requirements in all ACE programs in 2011–12, and having spent 85 percent or more of their fiscal year 2012 grant award).

Consistent with the requirements for ACE programs, the grantee agencies' requirements were to develop SPP programs aligned with the ACE program's Critical Success Model (see Appendix A). Specifically mentioned in the RFP was that:

- Innovative instructional techniques based on research and best practices be implemented
- Adult advocates support student involvement in school
- Preassessment and postassessment data be used to identify student needs and provide targeted interventions
- Professional development be provided to staff to increase their effectiveness

In November 2012, 15 21st CCLC grantee agencies were awarded SPP supplemental grants through a competitive process and established SPP programs in 46 centers where they also were implementing ACE programs. That means that in 2012–13, approximately 5 percent (46 out of the 864) of all Texas 21st CCLC centers had SPP programming, and approximately 39 percent (46 out of 119) of the centers run by SPP grantees had SPP programming. The grantee agencies included 10 school districts (Austin, Fort Worth Cycles 6 and 7, Harlingen, Richardson, Sherman, Snyder, Temple, Taylor, and Valley View); one regional education service center (ESC 13); two community-based organizations (Communities in Schools, one in San Antonio and the other in southeast Harris County); one county department of education (Harris County); and one charter school (NYOS). The SPP grants were reawarded in the 2013–14 school year, but two Cycle 5 grantees ended their programs at the conclusion of the 2012–13 school year. Therefore, only 13 grantees operating 94 centers, 37 of which had SPP funding, were included in the 2013–14 evaluation.

Statewide Evaluation: Year 1 (2010–11) and Year 2 (2011–12)

Beginning in fall 2010, TEA contracted with American Institutes for Research (AIR) and its partners at Gibson Consulting Group and the David P. Weikart Center for Youth Program Quality to conduct a statewide evaluation of the Texas ACE program, geared toward two primary research objectives TEA established for the project:

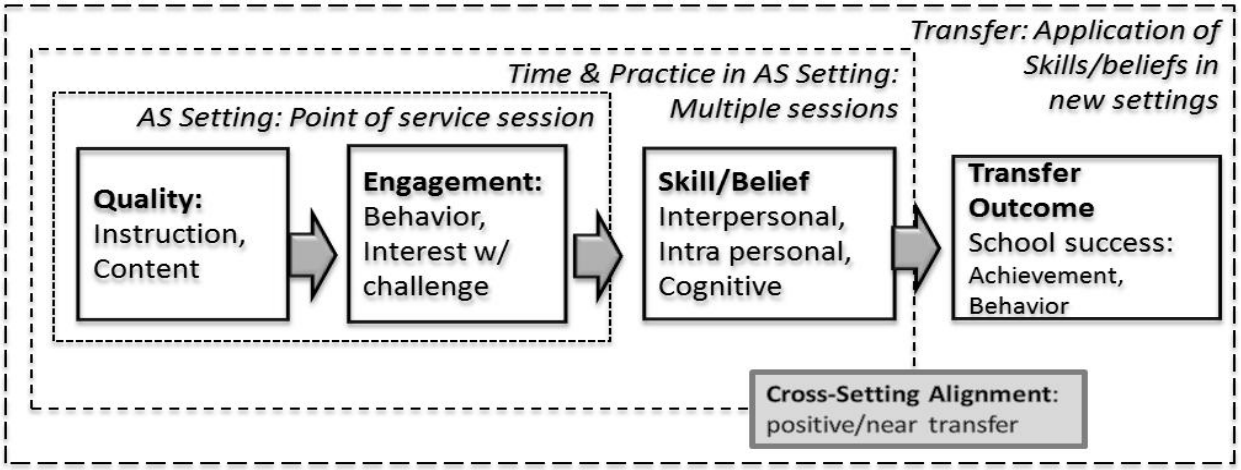
- **Research Objective 1:** Identify and describe innovative strategies and approaches implemented by successful 21st CCLC programs.
- **Research Objective 2:** Conduct a statewide assessment of 21st CCLC programs, operations, participation in the program, and student achievement outcomes.

For the first two years of the evaluation, AIR focused on answering the following evaluation questions associated with these two objectives:

- To what extent do students participating in services and activities funded by ACE demonstrate better achievement (along with other student outcomes) compared with similar students not participating in the program (separate estimates for which participation is defined as 1 day, 30 days, and 60 days)?
- To what extent do students who have higher participation rates (60 days or more) demonstrate better achievement (along with other student outcomes) compared with similar students who participate in ACE at lower levels (30 days to 59 days)?
- Does ACE's impact on student outcomes vary by relevant program characteristics, particularly program quality?
- What impact does participation in programming have on academic mindsets and behaviors, including *Self-Efficacy*, a *Mastery Orientation*, *Effort and Persistence*, and key learning behaviors?
- Are particular instructional and organizational practices and approaches more likely to produce positive changes in students' engagement?

The collective knowledge generated by answering this set of evaluation questions has allowed the evaluation team to explore several elements of the theory of child-level change outlined in Figure 1 and originally developed by the David P. Weikart Center for Youth Program Quality, including the ACE program's impact on transfer outcomes (i.e., TAKS or STAAR performance in reading and mathematics, school-day attendance, grade promotion); the relationship between program quality and transfer outcomes; and the relationship between program quality and student engagement. This theory of change has been used by multiple out-of-school time (OST) initiatives around the country and is not specific to the Texas ACE program. However, the concepts are applicable, and AIR has used this theory to guide the overall evaluation design.

Figure 1. Theory of Youth Change Related to Afterschool Program Participation



AS = afterschool

The results of the first two years of the evaluation are presented in two reports (Naftzger, Manzeske, Nistler, & Swanlund, 2012; Naftzger et al., 2013). Key findings from those two years were as follows:

- There were three instructional approaches found to be associated with high levels of student engagement: clarity of purpose, intentional use of time, and an active and interactive activity leader.
- Organizational practices associated with high levels of quality include intentional program design, staff development and collaboration, methods to monitor improvement, linkages to the school day, and community connections.
- There was some evidence of a connection between high-quality programs and high levels of participation.
- Higher levels of attendance (60 or more days) in 21st CCLC–funded programs were associated with higher levels of TAKS reading/ELA and mathematics performance, reduced disciplinary incidents and school-day absences, and supported grade promotion. In addition, a high level of point-of-service (POS) quality was associated with greater impact on reduced disciplinary incidents and grade promotion.
- For high school students, participation in an ACE program increased the likelihood of being promoted to the next grade level by 97 percent. There were similar findings for elementary and middle school students, but the magnitude was much smaller.

Statewide Evaluation: Year 3 (2012–13) and Year 4 (2013–14)

The introduction of the SPP program led to a change in the evaluation focus, although the scope remains the same. Beginning with the 2012–13 evaluation, AIR focused its activities and questions specifically on the SPP program as it compared with traditional ACE programming. The remainder of this report is dedicated to the findings from the 2012–13 and 2013–14 evaluation activities and analyses.

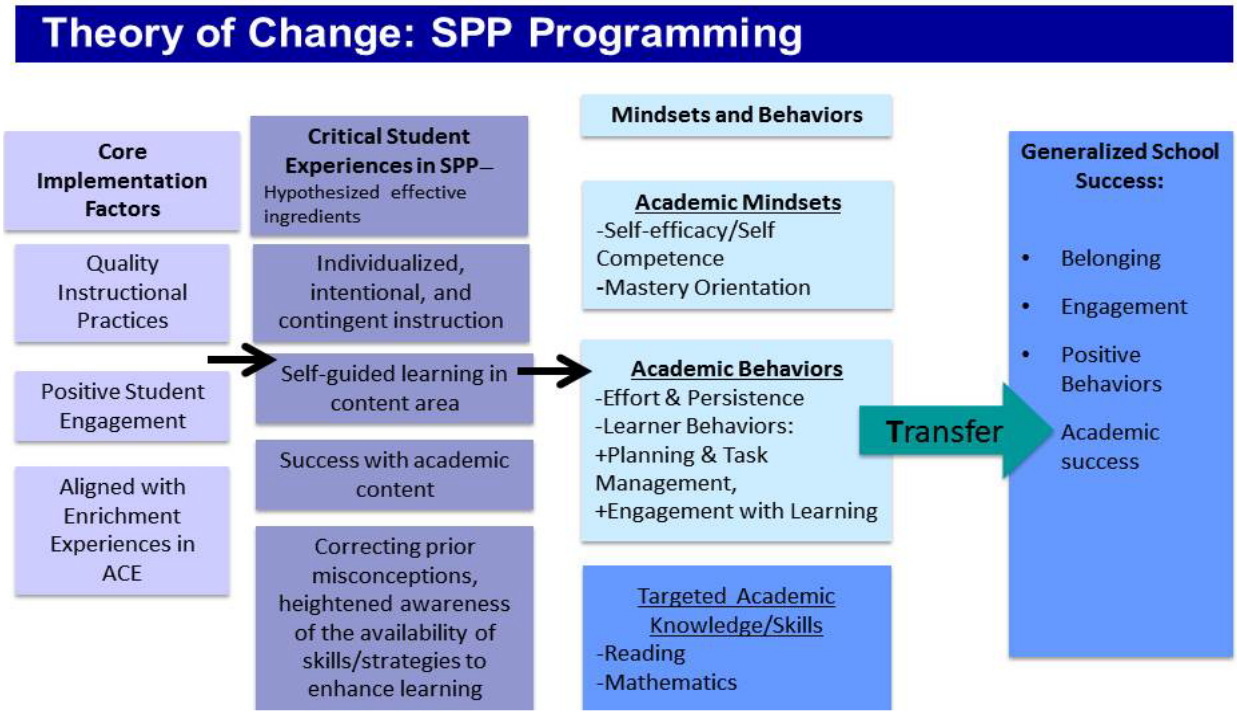
Linking Quality to Outcomes Through a Theory of Change

As noted previously, program and activity quality are expected to affect student engagement and, thus, student learning. This theory is supported by a 2007 meta-analysis examining the connection between developing personal and social skills in afterschool settings and a range of outcomes, including academic achievement (Durlak & Weissberg, 2007). The study found that afterschool programs employing what the authors dubbed the *S.A.F.E. features* (for sequenced, active, focused, and explicit) had significant benefits for youth on a wide range of outcomes. Programs missing these four features showed no significant outcomes. Since the release of this study, additional research has emerged on the importance of POS quality in OST and expanded learning programs. For example, a rigorous study conducted by Smith et al. (2012) showed that, by intentionally focusing on data-driven improvement, programs can and do become higher quality. Other research has found that programs with sufficient quality, breadth,

intensity, and duration can achieve a variety of social-emotional and academic outcomes (Auger, Pierce, & Vandell, 2013; Kataoka & Vandell, 2013; Kauh, 2011). More specifically, youth who report high levels of engagement in their afterschool programs have better prosocial skills, task persistence, and work habits (Kataoka & Vandell, 2013). This research supports the generic theory of change model outlined previously as well as a second theory of change that was developed by the evaluation team in partnership with TEA.

The SPP theory of change, depicted in Figure 2, articulates the key facets of SPP implementation that contribute to the experiences youth may have in the program. A sequence of high-quality, engaging experiences across time will lead to students developing key beliefs and skills, both social-emotional and academic in nature. Improving these key mindsets and behaviors will, in turn, affect youth performance on key metrics during the school day. The specifics of the theory of change follow Figure 2.

Figure 2. SPP Theory of Change



Core Implementation Factors

Quality Instructional Practices. As shown in Figure 2, high-quality instruction is one of the core implementation factors that leads to the second column of critical student experiences in SPP programming. Quality instructional practices are conceptualized as a series of practices and approaches that result in the creation of developmentally appropriate, high-quality settings for youth at the POS. These practices and approaches include the following:

- **Organizational-Level Practices.** Practices, structures, and approaches adopted by the **organization** as a whole, such as how staff are hired, oriented, trained, and evaluated; how partners are selected and integrated into program operations; and how the program creates developmentally appropriate settings aligned with core youth development principles and meaningful and relevant academic content
- **Activity Leader–Level Practices.** Practices and approaches directly adopted by **activity leaders** to make activity sessions supportive, interactive, engaging, and cognitively stimulating at the POS

Data related to program quality are important because they make it possible to explore the relationship between program quality and youth outcomes.

Positive Student Engagement. As the model posits, students need to engage positively in activities in order to benefit from participation. Thus, for the evaluation of the SPP programs, data on youth engagement are essential.

Aligned Enrichment Experiences. A third core implementation factor is aligned enrichment experiences that allow youth to engage in fun, hands-on learning that extends their SPP academic activities and allows for continued skill development.

Critical Student Experiences in SPP

The second column in Figure 2 indicates the critical experiences students must have in the SPP programming in order to develop the mindsets and behaviors as intended. These experiences represent high-quality practices, such as individualized instruction, self-guided learning, and correcting knowledge, so that students can have success with the content.

Mindsets and Behaviors

If youth sustain engagement in high-quality activities throughout multiple sessions, column 3 in Figure 2 hypothesizes that they will change their mindsets and behaviors as a result of their participation. These mindsets and behaviors include a *Sense of Self-Competence* and *Self-Efficacy*, an orientation toward mastery, an ability to engage with learning, and a willingness to persist. With the change in academically relevant mindsets and behaviors, specific, content-related skills in areas such as reading and mathematics may be strengthened, particularly as a result of individualized academic support to students.

Generalized School Success

Finally, as shown in Figure 2, the last column in the theory of change depicts learning transferred to the school setting. Context-specific mastery experiences support longer term skill development and skill transfer to the school day, ultimately leading to improved academic outcomes, such as improved grades and scores on state achievement tests, greater sense of attachment to school, and reduced dropout rates.

Research Questions

The evaluation of the SPP programs had six research questions (RQs) that continue to support the broader statewide evaluation objectives of impact and implementation and are designed to explore the validity of the SPP theory of change presented previously. Questions focus on describing the similarities and differences between ACE-only programs and SPP programs, particularly differences associated with program operations, activities, and participating students; examining the quality of SPP programming, variations in quality across programs, and the program operations that SPP grantees have implemented to support quality programming; and learning how and the extent to which the SPP programming impacts participating students. The six RQs that guided the 2012–13 and 2013–14 evaluation activities are as follows:

- **RQ 1:** How does SPP programming compare with ACE-only programming in centers administered by the same grantee?
- **RQ 2:** How do students participating in SPP programming differ from students who participate in ACE-only programming?
- **RQ 3:** How does the quality of delivery differ between SPP and ACE-only programming?
- **RQ 4:** What instructional and administrative practices lead to high student engagement?
- **RQ 5:** How engaged are young people in SPP and ACE-only programming? What is the relationship between quality of delivery and student engagement?
- **RQ 6:** What is the impact of SPP programming on students' academic mindsets and behaviors? How does this compare with ACE-only students?

Data Sources and Methods

A comprehensive set of data sources and methods have informed the findings presented in this report. Data that provide the basis for the report include data collected from the SPP program staff through interviews and surveys, student pre-post surveys measuring key mindsets and behaviors, data collected through observations of program activities and student engagement surveys, and Texas administrative and student data. The data sources and methods are described in more detail in Appendix B.

Limitations

There are several important limitations to keep in mind when reviewing the data and findings in this report. First, the report combines information from 2012–13 and 2013–14. However, the report should not be interpreted as a longitudinal study. Rather, the two years should be taken separately. The students and staff interviewed during each year may be different, and there was no attempt to correlate the two. In addition, two of the grantees ended their programs after the 2012–13 school year, so the sample of centers was different from one year to the next. The data collection methods and analyses

performed during each year varied slightly, and, therefore, even findings that are similar from one year to the next should be interpreted cautiously and as separate findings.

The 2012–13 findings are based on a very short implementation timeline. SPP programming began for the first time in January 2013, and this evaluation looked at the first few months of programming. Those findings should be viewed as preliminary in nature. Findings from the 2013–14 evaluation also should be viewed as exploratory. The evaluation team used primarily descriptive and correlational techniques in reviewing data but did not perform quasi-experimental techniques to review the data. In 2014–15, the team will use more rigorous methods to examine youth and staff survey data to determine program impact. For now, the findings in this combined report should be viewed as exploratory and as guiding the direction for additional analysis.

Organization of the Report

This report contains a description of findings and recommendations from both the 2012–13 and 2013–14 evaluations. The 2012–13 study looked at the initial implementation of the SPP program when it was first piloted in spring 2013. The 2013–14 study looked at the SPP program after one full year of implementation. Because both evaluations focus on the SPP program at different points in its evolution, TEA and the evaluation team determined that a combined report, sharing findings from both years, was the best way to discuss the implementation and initial impact of the program. The report is organized around the SPP theory of change described previously. Each chapter begins by highlighting the section of the theory being addressed and the associated RQs. The report concludes with a summary of findings and recommendations and a presentation of next steps for the 2014–15 evaluation. Finally, the report contains extensive appendices that provide more detail about the evaluation methods, site visits, and instrumentation.

The content of the remainder of the report is as follows:

- **Chapter 2: SPP Grantees, Centers, and Students.** This chapter has two major sections. One section addresses the characteristics of the grantees and the centers where SPP programming was provided, and the second section addresses the characteristics of the students served by the SPP programs. The chapter examines key differences between the characteristics and the participating students of SPP programs and ACE-only programs that are administered by grantees that received SPP funding.
- **Chapter 3: Quality at the Organizational Level.** Chapter 3 discusses program quality by examining the differences between SPP programs and ACE programs in centers with SPP funding on a set of quality indicators. The chapter summarizes information obtained on the activity leader surveys.
- **Chapter 4: Quality and Engagement at the POS.** The discussion of quality at the POS, or activity level, includes an examination of the different typologies represented in the SPP programs, identifying different program types on the basis of the delivery format and on the general learning approach. The chapter

then examines the quality of programs and levels of student engagement and presents correlational analyses of the connection between the two. Finally, the chapter concludes with a description of quality practices captured from observations and interviews with staff at different SPP and ACE-only programs.

- **Chapter 5: Initial Exploration of SPP Impact.** Chapter 5 examines youth outcomes as measured by an Academic Mindsets and Behaviors Survey, first looking at the reliability and validity of the survey instruments, then looking at baseline differences between SPP program participants and the participants in ACE-only programming at the SPP-funded sites. A third analysis descriptively examines the pre-post differences in outcomes for SPP and ACE-only program participants.
- **Chapter 6: Conclusions and Next Steps.** Chapter 6 summarizes key findings from this report and describes the implications for the remainder of the evaluation.
- **Afterward:** The afterward focuses on plans for the 2014–15 evaluation, including a description of preliminary work on evaluating the cost-effectiveness of the ACE-only and SPP programs and a high school–specific study.
- **References**
- **Appendices**

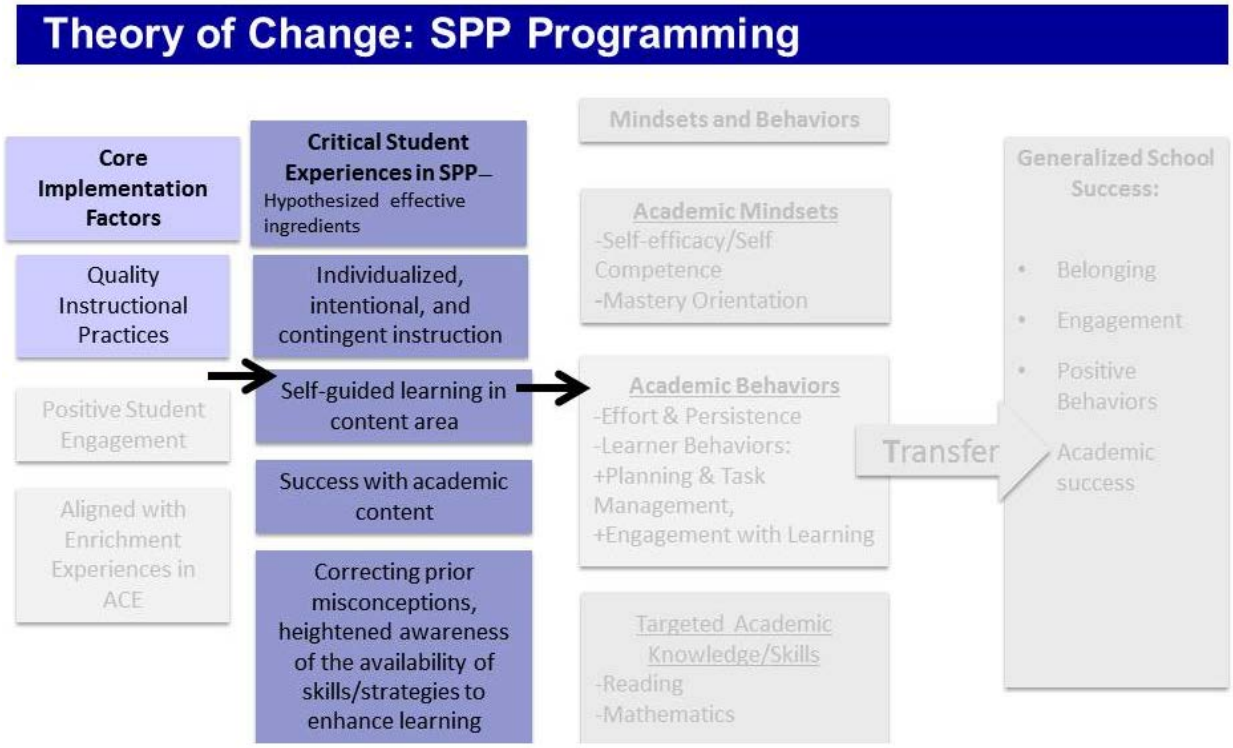
Chapter 2: SPP Grantees, Centers, and Students

SPP grants were awarded to 21st CCLC grantees to provide them with additional resources for offering academically focused support to students at risk for academic failure. Because of the Texas transition to a more rigorous academic assessment system (STAAR), TEA recognized that at-risk students may benefit from more academic support to achieve a satisfactory score on the assessment. The SPP grants were designed to address that need by tasking grantees to create programs with a more focused academic approach than ACE-only programs, programs that provide academic support but also nonacademic learning and enrichment opportunities. In fact, SPP programs were required to select an academic intervention that addresses local needs to implement as the focus of their programming.

Chapter 2 examines key features of the SPP programs related to grantee characteristics, program setting, staffing, and program activities. First, the chapter presents these features as descriptive and contextual information to help the reader understand what SPP programs looked like. Then, the chapter describes how SPP programs were different from ACE-only programs that were administered by the same grantees and how students who participated in SPP programs differed from the students participating in ACE-only programs at the same centers. This second portion of the chapter addresses both the core implementation feature of quality instructional practices as well as the Critical Student Experiences in SPP column of the SPP theory of change (see Figure 3) in order to provide context for the programs as well as answer two RQs.

- **RQ1:** How does SPP programming compare with ACE-only programming in centers administered by the same grantee?
- **RQ2:** How do students participating in SPP programming differ from students who participate in ACE-only programming?

Figure 3. SPP Theory of Change, Chapter 2



Grantee and Center Characteristics

The following sections provide an overview of the 15 ACE grantees that received the SPP funding; the centers the grantees operated; and, where possible, comparisons between ACE-only programs and SPP programs implemented by the grantees that received SPP supplemental grants. This information is useful for understanding the context in which SPP programs were implemented as well as how they compared with ACE-only programs.

Grantee Organizational Type

One of the features of the ACE program is that all types of organizations are eligible to apply for and receive 21st CCLC grant funding. In drawing distinctions between grantees, one of the biggest differences is in terms of whether a grantee is a school-based entity. One of the findings in the 2010–11 evaluation report (Naftzger et al., 2012) was that school-based grantees had more information about student academic needs, including student performance on standardized assessments, as well as school-day information about instruction and curriculum. Thus, school-based grantees were more likely to deliver programming that was aligned with the school day. Programs aligned with the school day may be better able to provide programming that supports the development of specific targeted academic skills and outcomes.

In 2012–13, 11 of the 15 grantee agencies with SPP funding (approximately 73 percent) were school based (according to TEA's TX21st CCLC Student Tracking System [TX21st]).[13] Among the 2012–13 21st CCLC grantee agencies overall, including grantees that did not receive SPP funding, 80 percent were school-based entities, indicating grantees awarded SPP funding were slightly less likely to be school based than were 21st CCLC grantees on the whole. In 2013–14, two of the SPP grantees had ended their funding cycle, one grantee being school based and one grantee being a nationally affiliated nonprofit agency. The percentage of grantees being school based remained almost the same (77 percent). Because SPP programs were intended to be more academically focused than ACE-only programs, the fact that SPP grantees were less likely to be school based is interesting and may mean those grantees that were not school based had a slightly harder time making their program well aligned with the school-day instruction.

Grantee Maturity and Cohort

As noted previously, eligibility for SPP funds was limited to 21st CCLC grantee agencies that were experienced in implementing ACE programming and demonstrated that they supported student participation and persistence by providing effective and engaging youth programming. As Table 1 shows, 2012–13 SPP funds were not awarded to any first-year 21st CCLC grantee agencies because the agencies had not yet proven they could provide effective programming and consistently comply with grantee requirements. In fact, as the table shows, the majority of SPP grantees received their 21st CCLC grant as part of Cycle 6, which was awarded in August 2009; 2012–13 was the fourth year of their ACE program administration, and 2013–14 was the fifth year of program administration. Two grantees, described as *sustaining*, were in their last year of 21st CCLC funding in 2012–13 and were not included in the 2013–14 data collection. This information is useful for understanding the context in which SPP programs were developed and delivered—to a large degree, SPP programs were implemented by grantees with a strong track record of performance and enough time left in their grant to pilot the program in 2012–13 before implementing it fully in the 2013–14 school year.

[13] TX21st is a database instrument used for tracking and reporting information about Texas 21st CCLC grant programs and the participants they serve. Fiscal agent grantee staff, as well as grant contacts and center coordinators, request access to the application in order to input participant and program data, run reports, and approve data for submission to TEA.

Table 1. SPP Grantees, by Maturity and Cycle (2012–14)

	SPP Grantees	
Grant Maturity	**Number of Grants**	**% Grants**
Mature	13	86.67%
Sustaining	2[1]	13.33%
21st CCLC Cycle		
Cycle 5 (2008–2013)	2[1]	13.33%
Cycle 6 (2009–2014)	9	60.00%
Cycle 7 (2011–2016)	4	26.67%
Total Grantees	15	100.0%
[1]The two sustaining Cycle 5 grantees were not included in the 2013–14 sample and data collection.		

SOURCE: TEA, TX21st, 2012–13 and 2013–14.

Center Hours

The ACE centers with SPP grants established hours of operation for their centers with SPP programming that were similar to the hours for their centers with ACE-only programs. According to TX21st, programming was offered approximately four days a week for an average of 16.5 hours per week for both ACE-only centers and centers with SPP programming. The majority of program hours in both types of centers (an average of 12 hours) took place after school. The balance of the hours took place before school. The only real difference between ACE-only centers and centers with SPP programming was the number of hours spent on average in morning programming. ACE-only centers ran programming in the morning, on average 3.1 hours per week, and centers with SPP programming ran programming in the morning, on average 2.4 hours per week. Overall, the hours of ACE-only centers and centers with SPP programming were not different, indicating that SPP programs were operating within the same confines of the ACE-only programming. This is likely because of the expectations of parents for pickup and scheduling as well as for staffing and logistical coordination.

Program Staffing

Survey responses from both SPP and ACE-only activity leaders, in both 2012–13 (*n* = 443) and 2013–14 (*n* = 395), indicated that, of the respondents, SPP activity leaders were about a third more likely to be certified teachers and to have a four-year college degree or higher than were ACE-only activity leaders.[14] As reported in the

[14] For 2012–13, activity leader survey responses indicated that 85 percent of the 246 SPP activity leaders who responded to the survey were certified teachers. In contrast, 51 percent of the 197 ACE-only activity leaders in centers with SPP funding who responded to the survey were certified teachers. Results for 2013–14 were similar. In that year, 81 percent of SPP activity leader respondents were certified teachers, and only 42 percent of ACE-only activity leader respondents were certified teachers. Moreover, survey results show that 88 percent of SPP activity leader respondents had a four-year college degree or higher, compared to 58 percent of ACE-only activity leader respondents.

2010–11 and 2011–12 evaluation reports, ACE program staff frequently included youth development workers and/or providers from the community who may lead enrichment activities and therefore are less likely to be certified teachers. These differences between SPP and ACE-only activity leaders are likely because of the emphasis of the SPP grant on providing academic support to at-risk students. Many site coordinators indicated in interviews that they specifically targeted certified teachers as activity leaders for SPP program activities because of their understanding of the subject material and facility with teaching specific content.

Program Activities

In addition to understanding the general characteristics of the SPP programs as they compare with ACE-only programs, it is also helpful to understand the nature of SPP activities compared with ACE-only activities in order to understand how their impact on participating youth may be different. Nationally, the goal of the 21st CCLC program is to provide academic and nonacademic enrichment programs that reinforce and complement the regular academic program of participating students. In the past, ACE programs have sought to meet that goal through a combination of generalized academic support (e.g., tutoring) and enrichment activities. The SPP grant sought to deliver on the goal more substantially by providing funding for centers to offer focused academic instruction in core subjects to a targeted group of students in need of further academic support in coordination with enrichment activities. As the theory of change posits (see Figure 3), the SPP program would achieve results through quality instructional practices, engaged students, and aligned enrichment experiences. This next section of the report explores the first two columns of the theory of change by examining the critical experiences students had in SPP programs and beginning to touch on instructional practices.

This more intentional and intensive academic instruction in SPP programming was evident from the site visits to 15 centers offering SPP programs in 2012–13 and to 13 centers offering SPP programs as well as a matched sample of ACE-only centers in 2013–14, representing each of the grantee agencies. Site visits included interviews with site coordinators and activity leaders as well as observations of both types of programming. Site visits were conducted shortly after the SPP centers were implemented in 2012–13 and after almost a full year of implementation in 2013–14. Differences between the two types of programming that emerged from observations and interviews (in both centers with SPP programming and ACE-only centers) included the following:

- **SPP activities placed more focus on mathematics and reading instruction.** Each of the centers with SPP programs offered reading/ELA-focused programming, mathematics-focused programming, or both. For grantees that focused on both subject areas, typically, SPP programming included two types of academic interventions, usually one for reading/ELA and another for mathematics. Some of the centers scheduled sessions so that the two subject areas were covered in the same day, and some so that subject areas were addressed on alternate days. Although many of the ACE-only centers offered

mathematics and reading enrichment activities, they were not usually as structured as those in the SPP program, which is not surprising, given the intended purpose of the SPP program.

- **SPP activities had smaller group sizes.** The most common difference—observed by site visitors and noted by activity leaders in interviews—was that there were fewer students in the SPP activities than in the ACE-only activities, which, according to activity leaders, allowed for more individualized and differentiated academic support and more teacher–student interactions. One reason may be that targeted students were selected for the SPP program on the basis of academic need and therefore represent a smaller pool. The SPP program, as a whole, also was designed to provide students with more individualized support, which is accomplished more easily in small-group sessions than in large group-sessions. In 2012–13, activity leader survey respondents from SPP centers reported that they served approximately 7 students per week, and respondents from ACE-only programs reported they served approximately 13 students per week. In 2013–14, activity leader survey respondents from SPP centers reported higher numbers overall, but the difference between SPP and ACE-only programs was the same—that is, SPP program activity leaders reported that they served approximately 13 students per week, and respondents from ACE-only programs reported they served approximately 19 students per week.
- **SPP activities were longer in duration.** On the basis of the observed activities in both years, SPP academic activities were typically longer than the ACE-only academic activities provided at the same centers. For example, academic activities in one of the SPP centers were 90 minutes in duration, compared with 60 minutes in the center's ACE-only program. In another SPP program, academic activities were a continuous two hours, compared with 45 minutes in the ACE academic activities. This finding indicates that students in SPP activities received more academic intervention than did students in ACE-only activities.
- **SPP activities provided more focused academic instruction.** According to observations and activity leader interviews, SPP activities tended to focus on specific academic content with the goal of improving student skills in a specific subject area to increase performance on grades and achievement tests. ACE-only academic support tended to be more general and designed to help students with school work. For example, in several centers that were visited, the ACE-only program's academic component included homework assistance and small-group tutoring on an as-needed or requested basis. Students without homework or tutoring needs typically were allowed to work on an academically focused game (e.g., crossword puzzle). During observations, some ACE-only programs included academic enrichment activities designed to support learning but indirectly teach core content, such as science club or a student newspaper activity. In contrast, SPP activities observed provided direct instruction in mathematics or literacy. In some cases, the content was determined by an assessment of an individual student's performance level. For example, several sites offered a computer-based reading program that allows students to read an

excerpt on the computer and progress through the course levels on the basis of skill level, with teacher review and guidance.

These findings are in keeping with the SPP theory of change—that is, smaller group size, longer academic activities, and more focused academic instruction are all important for providing students with critical student experiences in SPP, including more individualized instruction and opportunities for success with academic content. Furthermore, many of these practices described previously are elements of high-quality instruction (e.g., smaller group size). It appears that, during the course of the SPP grant, centers were able to implement activities in a way that was aligned with the SPP theory of change.

Another difference between SPP programs and ACE-only programs was the main type of activities in which students in each program participated. Using information from TX21st on student attendance in different types of activities, evaluators examined the differences in both 2012–13 and 2013–14. Although the percentages varied from one year to the next, in general, a comparison showed the following:

- **SPP students spent a majority of their time in academic activities.** In 2012–13, SPP students spent approximately 82 percent of their program time in academic enrichment activities and tutoring sessions. In 2013–14, that percentage went down to 66 percent. In both years, they spent little or no time in homework help sessions (less than 10 percent in each year). In comparison, students participating in ACE-only programs at centers with SPP funding spent approximately 45 percent of their time in academic, homework help, and tutoring sessions in 2012–13 and 55 percent of their time in these activities in 2013–14.
- **SPP students spent little time in recreational activities.** In 2012–13, the SPP students spent approximately 8 percent of their time in recreational activities. In 2013–14, that number increased to 25 percent. In both years, the ACE-only students spent approximately 33 percent of their time in recreational activities.
- **SPP students spent less time on youth leadership and career readiness.** SPP students spent less time than ACE-only students (at centers with SPP programs) in activities associated with youth leadership and career readiness. The SPP students spent 3 percent of their time in these activities in 2012–13 and 6 percent of their time on these activities in 2013–14. In contrast, ACE-only students spent approximately 10 percent to 11 percent of their time in these activities in both years.

These differences are notable but not surprising. Because SPP programs were more focused on academic support and the immediacy of meeting STAAR requirements, it is logical that students in those programs would spend more time on academic enrichment and less time on recreation and youth leadership. The fact that SPP students began to participate in more recreational, youth leadership, and career readiness programs in 2013–14 compared with in 2012–13 (i.e., the jump from 8 percent to 25 percent in recreational program participation) could be attributed to the fact that in 2012–13, very few programs had students participate in nonacademic enrichment (which includes recreation and youth leadership activities) despite the requirement that they should. The

programs were focused on starting their academic intervention in the first semester of SPP implementation and did not place emphasis on enrichment activities until fall 2013. In 2013–14, because they were in their second year of implementation, programs were better able to follow the requirement that students in SPP programs also participate in nonacademic enrichment and recreation activities. Therefore, many programs did not achieve the full model presented in the theory of change—which clearly includes the importance of aligned enrichment experiences—until the second year of program implementation.

Youth Characteristics

This part of the report attempts to answer RQ2: How do students participating in SPP programming differ from students who participate in ACE-only programming? Because SPP programs served students who already are at risk of academic failure, it is expected that the students participating in SPP activities were different from ACE-only program participants, particularly with respect to academic achievement. That assumption is explored in the following sections.

Recruitment

A distinguishing characteristic of the SPP program was its recruitment focused on academically at-risk students. Although ACE-only programs also seek to enroll students in need of academic support, they often open enrollment to students on the basis of other types of need, such as students with behavioral or attendance problems. In many cases, after initially targeting students in need, ACE-only programs move to open enrollment to fill remaining slots. To learn more about how students were recruited for participation in the SPP programs, project directors from the grantee agencies were administered a brief recruitment survey in each year of the evaluation. This survey asked about their criteria for recruitment. The survey responses indicated the following:

- Initially, the intention was that students participating in the SPP program would be recruited primarily from outside of the ACE program in order to reach students most at risk academically who were not already attending ACE activities. In reality, students were recruited from both inside and outside the ACE program. In 2013–14, the second year of SPP implementation, most (85 percent) grantees reported recruiting students both from within and outside of the ACE program, and two grantee agencies recruited only students from within the ACE program.
- Most (93 percent in 2012–13 and 85 percent in 2013–14) SPP centers did not select students on the basis of demographic variables, such as their special education, economic, or English language learner status. In 2012–13, one of the SPP grantees recruited English language learner students. In 2013–14, one grantee recruited English language learner students, and one recruited students with free or reduced-price lunch status.

Students with high academic needs, meaning students at risk for academic failure, were identified systematically and selected for participation in SPP programs on the basis of local needs. This was a requirement of grantee applicants and was stated explicitly in

TEA's RFP. However, each center could define its own criteria for students' being at risk. According to the recruitment survey, as well as interviews with site coordinators, students were identified for participation by using multiple measures, including the following:

- Academic performance on the 2012 STAAR (primarily in mathematics and reading; only a few centers indicated they were using the science tests) or interim assessments (e.g., Istation's Indicators of Progress test, Dynamic Indicators of Basic Early Literacy Skills test)
- One or more additional academic performance measures (e.g., failing course grades)
- Recommendations by a teacher or a team of teachers, administrators, reading or mathematics specialists, and/or counselors

Again, centers were able to determine their own criteria for defining students' being at risk, which means there was a great deal of variation in the academic performance level of students who participated. Some centers were very specific about their criteria. For example, one SPP project director stated on the recruitment survey that the SPP programs would serve students who scored between 50 percent and 65 percent on the 2012 STAAR exam. Another SPP grantee selected students who scored below 50 percent on the STAAR reading exam. Others centers were less precise.[15] One project director said they recruited students who "scored at the bottom" in mathematics, reading, and science on the STAAR exam, although no specific threshold for enrollment in the program was indicated. One exception to recruiting the lowest performing students was an SPP grantee that selected "bubble" students—those near satisfactory on state assessments. This strategy was designed to help more students achieve satisfactory performance, thus raising the school's overall achievement level as well as increasing individual student performance.

Grade Level of SPP and ACE-Only Participants

Understanding the grade levels served by SPP and ACE-only programs can be useful for understanding the distribution of resources and where the intervention is most concentrated (Table 2). A review of the enrollment at centers operated by the 15 SPP grantees shows that, in spring 2013, 2,795 students were enrolled in SPP activities at ACE centers with SPP funding. In addition, 6,105 students were enrolled in ACE-only programming at those centers. That means there was a total of 8,900 students enrolled in programming at these centers, and, thus, 31 percent of the ACE student population at SPP grantee centers attended SPP programming during the pilot of the SPP program in spring 2013. During the 2013–14 school year, 4,017 students were enrolled in SPP activities, and 7,190 students were enrolled in ACE-only programming, for a total of 11,207 students enrolled in ACE centers with SPP activities. Thus, approximately 36 percent of ACE participants at SPP grantee centers attended SPP programming during

[15] Because STAAR was new in 2011–12, there was a delay in setting the passing standard for the assessment. This delay may have led some centers to be precise in their recruitment criteria for the 2012–13 school year and some centers to opt for less precise measures.

the 2013–14 school year. These enrollment numbers and percentages should not be compared because in 2012–13 the analysis looks at students enrolled in programming just during the spring of 2013 because SPP programming was not implemented until then. The 2013–14 analysis looks at students enrolled in both types of programming for the entire school year because the SPP program was implemented for a full year. It would be expected that enrollment would be higher for a full year because new students can enroll each semester. Furthermore, some centers overenrolled students through their recruitment process, so actual participation rates may be lower than these enrollment figures would indicate. Finally, the enrollment figures in Table 2 are mutually exclusive across years, which means they may represent different groups of students.

Table 2. Grade Levels of SPP and ACE-Only Program Enrollees (2012–14)

Grade Level	2012–13				2013–14			
	SPP Students[1]		ACE-Only Students at Centers With SPP Funding[1]		SPP Students[1]		ACE-Only Students at Centers With SPP Funding[1]	
	N	%	N	%	N	%	N	%
K	31	1.1%	54	0.9%	45	1.1%	285	4.0%
1	134	4.7%	254	4.2%	125	3.1%	671	9.3%
2	197	7.0%	498	8.2%	200	5.0%	804	11.2%
3	500	17.7%	833	13.6%	547	13.6%	692	9.6%
4	527	18.6%	911	14.9%	519	12.9%	814	11.3%
5	282	10.0%	654	10.7%	252	6.3%	901	12.5%
6	198	7.0%	401	6.6%	148	3.7%	393	5.5%
7	322	11.4%	658	10.8%	351	8.7%	490	6.8%
8	205	7.3%	437	7.2%	311	7.7%	400	5.6%
9	181	6.4%	481	7.9%	399	9.9%	248	3.5%
10	203	7.2%	534	8.7%	538	13.4%	510	7.1%
11	13	0.5%	260	4.3%	436	10.9%	473	6.6%
12	2	0.1%	130	2.1%	146	3.6%	480	6.7%
Total	2,795	99%[2]	6,105	100.1%[2]	4,017	99.9%[2]	7,190	99.7%[2]

[1] Enrollment data are mutually exclusive. In 2012–13, the analysis looks at students enrolled in programming just during spring 2013. The 2013–14 analysis looks at students enrolled in both types of programming for the entire school year.

[2]Percentages may not equal 100 percent because of rounding.

SOURCE: TEA, TX21st, 2012–13 and 2013–14.

As Table 2 shows, the percentages of SPP and ACE-only students at centers with SPP funding was similar across years. Enrollment in the program seems to be most concentrated (more than 10 percent of enrolled students) in Grades 3, 4, and 7 in 2012–13 and in Grades 3, 4, 10, and 11 in 2013–14. The emphasis on Grades 3, 4, and 7

may be because, in Grade 5 and again in Grade 8, students who do not pass STAAR are required to receive accelerated instruction and are subject to retention. Supporting Grade 3 and 4 students may be a strategy to reduce the numbers of students not passing STAAR in Grade 5. In 2012–13, very few Grade 11 and 12 students enrolled in SPP programming, although in 2013–14 quite a large number did. It is unclear from the recruitment survey and interviews why participation of Grade 11 and 12 students would be so different from year to year. It may be that students who were in Grade 10 in 2012–13 continued participation in 2013–14 when they were in Grade 11. This situation, however, accounts for only some of the difference. The difference also may reflect changing priorities for academic improvement in certain centers and schools or simply inflated enrollment numbers in some high school centers.

Academic Performance

As noted previously, a key and expected difference between the SPP students and the ACE-only students at centers with SPP funding is related to academic performance. SPP programs were designed to target students in need of more intense academic support. To examine whether centers with SPP programming were successful in recruiting students at the outset of programming in spring 2013 who were at risk for academic failure, the evaluation team examined the 2012 STAAR exam results for participating SPP students. This examination showed that SPP students failed to meet the phase-in 1 Level II (Satisfactory) standard at a higher rate than did ACE-only students.[16]

Table 3 shows the number and percentage of students who met different standards on the 2012 STAAR exam. Because the STAAR exam was more rigorous than its predecessor (TAKS), in 2012, TEA created a phase-in 1 Level II (Satisfactory) standard and identified a threshold for satisfactory performance that was somewhat lower than the STAAR recommended standard would be after the phase-in period. At the phase-in 1 Level II (Satisfactory) threshold, fewer than 50 percent of students who eventually would participate in SPP programming had met the standard in reading or mathematics (for elementary and middle school students). For the end-of-course assessments for Algebra I and English I-Reading (high school students), only 43 percent and 25 percent of SPP students, respectively, met the phase-in 1 Level II (Satisfactory) standard. This percentage was lower than for the ACE-only students—particularly on the English I-Reading assessment.

[16] At the time of this analysis, the standards had been set for the 2011–12 STAAR assessment.

Table 3. 2011–12 Results on STAAR for 2012–13 SPP and ACE-Only Students for Phase-In 1 Level II (Satisfactory) Standard

	SPP Students		ACE-Only Students	
STAAR 2011–12[1]	**N[2]**	**%**	**N[2]**	**%**
Grades 3–8 Mathematics	815	48.95%	2,313	67.61%
Grades 3–8 Reading	785	47.15%	2,420	70.74%
EOC[3]: Algebra I	91	42.92%	373	75.66%
EOC[3]: English I-Reading	52	24.53%	304	60.08%

[1] The data presented here are the percentage of students in each category across all grade levels tested.

[2] *N*s on the academic performance tables may vary from those on the enrollment tables because not all students at all grade levels are testing for each exam. For example, the STAAR mathematics exam is administered to students only in Grades 3–8, so only students in the appropriate grade levels would be counted in these categories in the table.

[3] EOC refers to end-of-course assessments, in which students taking the courses described were assessed at the end of the year.

SOURCE: TEA, STAAR exam, 2011–12.

The STAAR test results highlight that the intention of the SPP program—to recruit and serve academically at-risk students—was borne out in its first year of implementation. Not only did SPP students present as below proficiency on state assessments, but they were less likely to have achieved the phase-in 1 Level II (Satisfactory) standard for academic performance than were their peers in ACE-only programs, again validating that the intended goal of the SPP program was being met.

Conclusions and Key Findings

This chapter focuses on several elements of the theory of change—quality instructional practices and critical student experiences in SPP, such as individualized instruction, alignment with enrichment, and opportunities for success with academic content. The goal for this chapter was to examine the characteristics of SPP programs and their students compared with those of ACE-only programs and students within the same center in order to understand better the context for these programs and the differences between the two. The chapter attempted to answer two RQs:

- **RQ1:** How does SPP programming compare with ACE-only programming in centers administered by the same grantee?
- **RQ2:** How do students participating in SPP programming differ from students who participate in ACE-only programming?

The answers to these questions will be useful to frame the discussion of program quality at the POS and the organizational level as well as to examine the initial impact of SPP programs on participating students. The key takeaway from this chapter is that the characteristics of SPP programs and SPP students differed quite substantially from ACE-

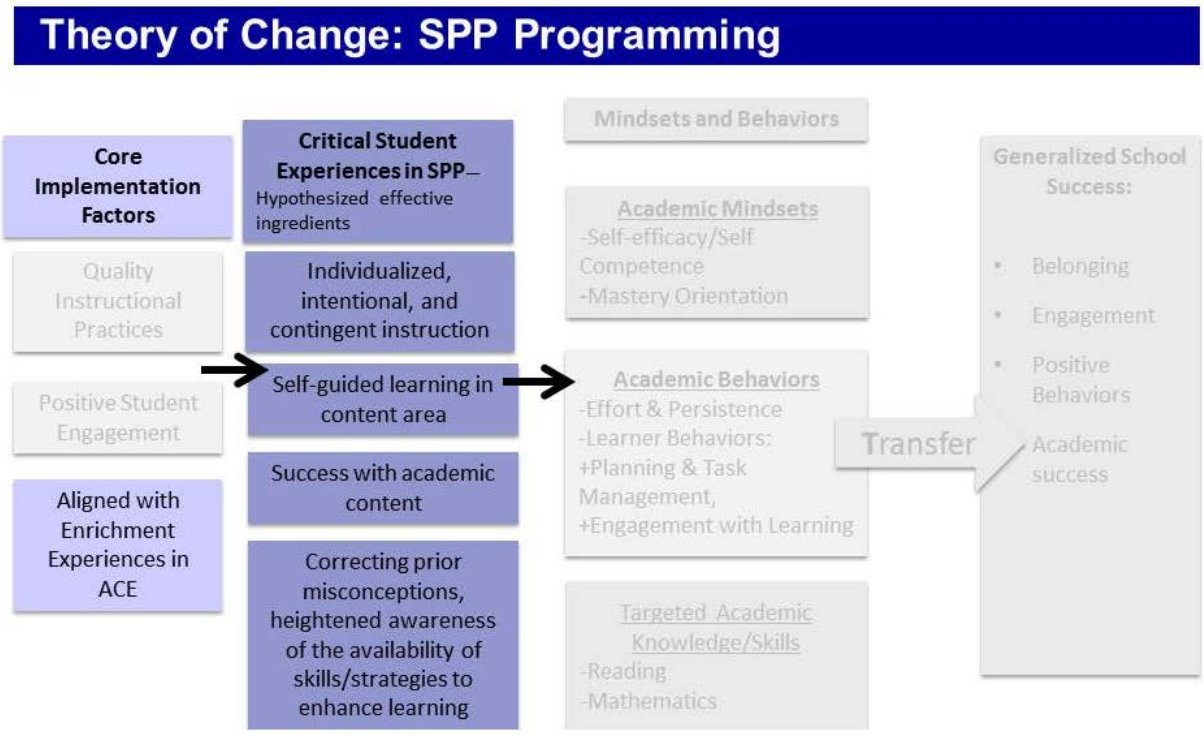
only programs and students, despite being within the same center. More specifically, analyses revealed the following:

- Centers with SPP programming were experienced centers with similar center operations (e.g., hours and grade levels) as ACE-only programs.
- SPP programs differed in their staffing and activities from ACE-only programs, showing a greater academic orientation in the following ways:
    - **SPP activities were staffed primarily by certified teachers** (about 85 percent of the time), and ACE-only programs were staffed by a mix of certified and noncertified teachers (about 50 percent of each). This may be because SPP programs are more focused on instruction in core subjects (e.g., mathematics and literacy) and certified teachers were targeted specifically to staff the program.
    - **SPP students spent more of their time in academic activities than ACE-only students**—particularly in academic enrichment activities and tutoring. Unlike ACE-only students, SPP students spent very little time in recreation, youth leadership, and career readiness activities. Although this difference was less substantial in 2013–14, there remained a difference between the amount of time SPP students and ACE-only students spent in academic versus recreational or youth leadership programs.
    - **SPP students typically met in relatively small groups**, and SPP activity leaders served fewer students per week than did ACE-only activity leaders. According to activity leaders, the smaller groups allowed for more intensive one-on-one support and responsiveness to individual student needs.
    - **SPP and ACE-only students were similar in terms of grade level but differed in terms of academic performance.** Students who attended SPP programs had higher academic needs than did ACE-only students, as shown by the lower percentage of SPP students who achieved satisfactory performance levels on the STAAR assessments. Although recruiting at-risk students was the stated goal of the SPP grantees, analysis of the data highlights the fact that the SPP programs were recruiting at-risk students, as intended, and were serving students with higher academic needs than were ACE-only programs.

Chapter 3: Quality at the Organizational Level

Quality at the organizational level refers to the organizational and instructional practices that staff members enact to provide coherence and consistency for activities. An organizational climate that supports staff in reflecting on and continually improving program quality is a key aspect of effective youth development programs (Birmingham, Pechman, Russell, & Mielke 2005; Glisson, 2007; Smith, 2007). This chapter focuses on elements of the first two columns in the theory of change—core implementation factors and critical student experiences in SPP, including aligned enrichment experiences, individualized instruction, and opportunities for success with academic content that are supported and facilitated by the organization in a broad way. Chapter 4 covers some of these same program components but focuses on how they are supported by individuals at the POS. It is critical that staff in a program work to create an environment and set of experiences that are high quality and promote the goals of the program—through both high-quality implementation and creation of positive experiences for students.

Figure 4. SPP Theory of Change, Chapter 3



By examining the extent to which SPP programs, when compared with their more mature ACE-only program counterparts, had established practices and procedures that supported quality programming at the organizational level, Chapter 3 aims to address three RQs partially, as follows:

- **RQ1:** How does SPP programming compare with ACE-only programming in centers administered by the same grantee?

- **RQ3:** How does the quality of delivery differ between SPP and ACE-only programming?
- **RQ4:** What instructional and administrative practices lead to high student engagement?

This chapter addresses these questions only partially. For example, the chapter looks at high-quality administrative practices, which are important for the overall quality of program delivery. The chapter examines how administrative practices (a key aspect of programming) differ between ACE-only and SPP programming and closely examines administrative practices that are in use, but the chapter does not yet connect those to high student engagement. That analysis is described in Chapter 4 as is the more detailed analysis of instructional practices.

Organizational Quality: Organizational Practices

This section focuses on organizational practices that contribute to the quality of the overall program. The data that inform quality at the organizational level are responses to the activity leader survey, which was administered in both years of the evaluation to a sample of leaders of SPP activities as well as leaders of ACE-only activities within the same center. The evaluation team conducted analyses of survey data to determine whether the quality of programming was different in SPP and ACE-only activities (see Appendix C for detailed survey responses broken down by question and staff type), thereby addressing RQ1: How does SPP programming compare with ACE-only programming in centers administered by the same grantee? and RQ3: How does the quality of delivery differ between SPP and ACE-only programming? As noted in the previous chapter, SPP programs provided different content (more academic support versus a combination of academics, enrichment, and recreation) and often were led by certified teachers rather than youth workers or other nonteaching staff. On the one hand, these differences may have led to differences in program quality. On the other hand, SPP programs may have established organizational practices that were similar to those of the ACE-only programs because they were directed by the same grantee agencies and were located in existing ACE centers.

To examine organizational quality, the evaluation team applied a statistical technique called *Rasch analysis* on the staff survey responses (see Appendix D for details on the analysis methods) for two specific constructs: *Internal Communication* and *Program Climate.* A construct is a set of items, or questions, on a survey that are all related to the same theme or concept. For *Internal Communication*, the survey contains eight items that, together, tell evaluators the extent to which program staff have different opportunities to discuss, plan, and improve programming. Likewise, the *Program Climate* construct is made up of nine items that, together, demonstrate the extent to which there is a cooperative environment that allows continuous improvement efforts to take hold. Rasch analysis produces a scale score that puts all survey responses on a scale from 1 to 100 for easier analysis and to ensure all items are measured the same way. The items that were included for each of these two constructs from the activity leader survey are presented in Appendix E, Table E-1.

The analysis of the responses for these two constructs indicated that, overall, there were relatively high degrees of agreement among both types of activity leaders on *Internal Communication* and *Program Climate*—that is, staff members were likely to indicate that communication practices occurred often or that they agreed or strongly agreed that certain practices related to the two constructs were present. The high levels of communication indicate that program quality was present on these organizational practices.

The analyses also showed that, although both groups reported agreement, there were differences between SPP and ACE-only activity leaders in terms of their perceptions of *Internal Communication* and *Program Climate* (see Table 4). During the first year of SPP program implementation (2012–13), the differences were most pronounced on the *Internal Communication* construct. This construct includes topics such as communicating about student needs, having conversations about improvement, and participating in joint professional development. When the evaluation team looked at scale scores for the *Internal Communication* construct, the difference between the two types of instructors was moderately significant,[17] which means that the findings can be believed with a moderate degree of confidence. In this case, the moderate significance indicates that SPP activity leaders engaged in strategies associated with *Internal Communication* (i.e., meeting once a week) more frequently than did ACE-only activity leaders. In 2013–14, the difference between the two groups was smaller and no longer statistically significant.

Although SPP activity leaders reported engaging in *Internal Communication* strategies more often than did ACE-only activity leaders, both types of activity leaders reported engaging in the same types of strategies. For example, in 2012–13, for both ACE-only and SPP activity leaders, the method of *Internal Communication* used most frequently was *Collaborate on Ways to Improve Student Engagement*, and the method used least frequently was *Use Program Data for Program Planning.* In 2013–14, the method of *Internal Communication* cited most frequently was *Work Together to Share Ideas and Approaches to Effectively Meet Individual Student Needs* (63 percent of staff said they did this weekly), and the least frequent method was *Use Program Data for Program Planning* (37 percent said they did this weekly). See Appendix C, Table C-5 for more detail on the 2013–14 results.

SPP and ACE-only activity leaders in both 2012–13 and 2013–14 appeared to feel differently from one another in terms of *Program Climate* as well as *Internal Communication.* This time, however, it was the ACE-only activity leaders who were statistically more likely than SPP activity leaders to agree with the items related to the *Program Climate*, as shown in Table 4. These items include topics such as enjoying

[17] Statistically significant differences are defined in this report as those with a *p* value of 0.05 or below, which means that there is 5 percent probability (or less) of randomly observing a difference of this size or greater if no difference exists. A moderately significant difference would yield a *p* value less than 0.1, where there would be a 10 percent probability of observing a difference of this size by chance. Essentially, these two terms help to illustrate the degree of confidence the research team has in the findings.

their job, having high morale, having enough time to prepare, and feeling a connection to program management.

Table 4. Scale Scores for *Internal Communication* and *Program Climate* (2012–13 and 2013–14)

	Construct	SPP Mean[1]	ACE Mean[1]
2012–13	*Internal Communication**	50.78 (Nearly Every Week) *N* = 242	49.00 (About Once a Month) *N* = 188
	*Program Climate**	49.24 (Agree) *N* = 245	50.96 (Agree) *N* = 193
2013–14[2]	*Internal Communication*	50.61 (Nearly Every Week) *N* = 208	49.27 (About Once a Month) *N* = 173
	*Program Climate***	48.93 (Agree) *N* = 211	51.22 (Strongly Agree) *N* = 184

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

[1] Numbers appear similar from one column to another but have different response categories because of the statistical technique used in analysis that assigns respondents a raw score on a scale of 1 to 100. Raw scores then are transformed into scale scores on the basis of the mean response for all survey participants that then are correlated to the response categories on the survey. The transformed scale may be very close to the cutoff point for a particular response category, so scores appear close, but response categories differ.

[2] Because staff members surveyed and interviewed each year are different, it is important not to compare 2012–13 data with 2013–14 data.

SOURCE: 21st CCLC Evaluation Activity Leader Surveys, 2012–13 and 2013–14.

Although the answers varied slightly from year to year, it appears, from the responses, that SPP activity leaders may have had a higher degree of *Internal Communication*, and ACE-only activity leaders may have had a better perception of *Program Climate*. Although the reasons for this finding cannot be determined with certainty by looking at survey responses, there are several factors that may have contributed to this finding. Because 2012–13 was the first year of SPP implementation, it is possible staff had to do more planning and meeting to ensure the program was implemented as planned. It is also worth noting the SPP activity leaders reported meeting more frequently in both years. Because the programs tend to involve several activity leaders implementing the same program within a center and even across centers within one grantee, there may have been more professional development and joint planning than among activity leaders teaching disparate types of activities (e.g., homework help and sports). SPP activity leaders may have reported more often that they do not have enough time and resources to address individual student needs because they are tasked with improving academic achievement and may feel that task requires more intensive support than they can provide in a short afterschool session. SPP programs also target students most in

need of academic intervention, so SPP activity leaders may have felt they did not have adequate time to address the individual needs of all students in their programs. Overall, these findings tell us that ACE-only and SPP activity leaders differed but also that they had a high degree of agreement on the two constructs, which indicates that program quality as measured by these constructs was present for both types of instructors in different ways.

Organizational Quality: Activity Leader Practices

This section discusses organizational practices in which activity leaders engage that contribute to high-quality instruction. These practices support instructional practices discussed in the next chapter and are part of the second column in the SPP theory of change related to critical experiences for students in SPP programming, such as intentional instruction. This section begins to address RQ4: What instructional and administrative practices lead to high student engagement? Although this chapter does not yet make the connection to engagement, it examines administrative practices that may contribute to or take away from student engagement. The data for this section are also taken from staff surveys.

Planning

Because SPP programming is oriented toward providing at-risk students with appropriate academic support, the expectation is that SPP program activity leaders, more than ACE-only program activity leaders, plan activities with an orientation toward identifying and addressing student learning goals. To understand better the difference between how the two types of activity leaders engage in planning, the evaluation team examined two constructs—*Access to Student Data* and *Intentional and Responsive Instruction*. *Access to Student Data* items refer to both the availability of data and the use of student academic data, such as STAAR assessment scores and grades. *Intentional and Responsive Instruction* items address individual planning of lessons as well as planning with other staff and school staff—particularly with respect to establishing learning goals and the emphasis on skill building (see Appendix E, Table E-2 for a list of items making up these two constructs and Appendix C, Tables C-1 and C-3 for response data on the survey items). Taken together, these two constructs paint a picture of how often and in what way activity leaders plan. Research has shown that programs that are focused and explicit (i.e., intentional) are more likely to produce desired outcomes than are programs that do not have those features (Durlak & Weissberg, 2007). Using data about student needs is one way to create more intentional programming that is focused explicitly on the needs of the student.

As Table 5 shows, in both 2012–13 and 2013–14, SPP activity leaders reported more regular adoption of behaviors related to *Intentional and Responsive Instruction* and *Access to Student Data* than did ACE-only activity leaders.

For both 2012–13 and 2013–14, SPP activity leaders had higher scores on the *Intentional and Responsive Instruction* construct than did ACE-only activity leaders. In both groups, the staff reported that they engaged in the practices frequently, but SPP

activity leaders reported, on average, that the practices *always occur*, but ACE-only staff reported that the practices *occur regularly.* The difference in means between SPP and ACE-only activity leaders was statistically significant, indicating that SPP staff were more likely to report adopting practices such as planning with other staff, using lesson plans, and planning in advance regularly than were ACE-only staff. On the *Access to Student Data* scale, although both groups, on average, reported that they *occasionally use* various types of student data, the difference between the two means was statistically significant, indicating that SPP activity leaders were slightly more likely than ACE-only activity leaders to have access to and to use student academic data. This finding was true for both years of the study.

Despite the similarity in results for 2012–13 and 2013–14, there were different staff members in place during each year. Although we can say that the results are similar from year to year, the results should not be viewed as longitudinal data.

Table 5. Mean Scale Scores on the Planning Constructs (2012–13 and 2013–14)

	Construct	SPP Mean[1]	ACE Mean[1]
2012–13	*Access to Student Data***	50.96 (Occasionally Use) *n* = 238	48.56 (Occasionally Use) *n* = 159
	*Intentional and Responsive Instruction***	50.99 (Always Occurs) *n* = 246	48.74 (Occurs Regularly) *n* = 193
2013–14[2]	*Access to Student Data****	51.54 (Occasionally Use) *n* = 209	47.95 (Occasionally Use) *n* = 157
	*Intentional and Responsive Instruction****	51.35 (Always Occurs) *n* = 211	48.48 (Occurs Regularly) *n* = 188

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

[1] Numbers appear similar from one column to another but have different response categories because of the statistical technique used in analysis that assigns respondents a raw score on a scale of 1 to 100. Raw scores then are transformed into scale scores on the basis of the mean response for all survey participants that then are correlated to the response categories on the survey. The transformed scale may be very close to the cutoff point for a particular response category, so scores appear close, but response categories differ.

[2] Because staff members surveyed and interviewed each year are different, it is important not to compare 2012–13 data with 2013–14 data.

SOURCE: 21st CCLC Evaluation Activity Leader Surveys, 2012–13 and 2013–14.

Responses on individual items from the staff survey may help clarify the overall findings displayed in Table 5. For individual items (see Appendix E, Table E-2, for specific items associated with the two planning constructs and Appendix C, Table C-3, for specific response data for each item on the 2013–14 survey) on the *Access to Student Data* construct, in 2012–13 and 2013–14, both SPP and ACE-only activity leaders most often reported obtaining and using information from two sources: *Input From Students'*

School-Day Teachers (43 percent *often use*) and *Students' Grades* (42 percent *often use*). During that same time, both types of instructors were least likely to have access to or use *Results on State Assessments* (44 percent *do not receive this data*).

On the *Intentional and Responsive Instruction* construct, SPP activity leaders in 2012–13 were more likely than ACE-only activity leaders to report that they *Planned Instruction to Promote Skill Building and Mastery* in relation to state and district standards. In 2013–14, SPP activity leaders were more likely than ACE-only activity leaders to report that their activities were *Designed to Support Specific, Measurable Learning Goals for Students* (71 percent versus 56 percent of respondents, respectively, said this *always occurs*); were *Meant to Extend and Enhance Skills Developed in Prior Activity Sessions or Lessons* (67 percent versus 61 percent of respondents, respectively, said this *always occurs*); and were *Planned in Advance* (73 percent versus 62 percent, respectively, said this *always occurs*). ACE-only activity leaders were most likely to have *Planned in Advance* (62 percent). See Appendix C, Table C-1 for individual response data on the 2013–14 survey.

The responses on these constructs are, again, not surprising. Because SPP activities were designed to be more intensively focused on academics and specifically target students at risk for academic failure, it stands to reason that activity leaders for those programs would be more likely to use data to inform their programming and plan instruction focused on mastery of state standards. ACE-only activity leaders still reported high levels of *Intentional and Responsive Instruction*—they just reported it slightly less than did SPP activity leaders. Therefore, these data are not implying that ACE-only activities are not intentional and based on data but rather that SPP programs may have been even more intentional and data driven.

Instructional Constructs

Two of the survey constructs, *Individualized and Contingent Instruction* and *Academic Learning Activities*, address what the activity leader does in the activity in terms of responding to student needs and directing student learning activities. When instruction is individualized and contingent, it is based on student needs and directed toward helping a student meet specific learning objectives. *Individualized and Contingent Instruction* requires a teacher to know what kind of support a student needs, to know the student well enough to support him or her effectively, and to link instructional strategies with student needs. The SPP theory of change hypothesizes participating in activities using this type of instruction will contribute to better youth outcomes.

Also, because SPP programming is designed to improve student performance on the STAAR assessment, academic sessions were the main focus of SPP activities. Therefore, it is anticipated that certain *Academic Learning Activities* were more prevalent in SPP activities. To examine whether and how often these critical practices were taking place in the two types of programs, the evaluation team analyzed responses to survey items associated with the *Individualized and Contingent Instruction* and *Academic Learning Activities* constructs.

Analysis showed that SPP activity leaders in both years were slightly more likely than ACE-only activity leaders to respond *I Know This for Most of My Students* on the items related to *Individualized and Contingent Instruction* (see Table 6). These items included topics such as knowing a student's assessment scores, where the students are struggling in school, and their strengths and weaknesses. Although both groups of activity leaders responded, on average, that they knew this for most of their students, the differences between the two groups' mean scale scores was statistically significant, indicating that SPP activity leaders were slightly more likely to know this information about their students.

Also, in both years, there was a significant difference observed between SPP activity leaders and ACE-only activity leaders on practices related to *Academic Learning Activities*. These practices include topics such as allowing students to work independently and reviewing work from the school day. The SPP activity leaders reported, on average, that these practices *occur regularly*, but ACE-only staff members reported, on average, that these practices *occur occasionally*.

Table 6. Mean Scale Scores on the Instructional Constructs (2012–13 and 2013–14)

	Construct	SPP Mean[1]	ACE Mean[1]
2012–13	*Individualized and Contingent Instruction***	51.21 (I Know This for Most of My Students) *N* = 246	48.43 (I Know This for Most of My Students) *N* = 190
	*Academic Learning Activities****	51.74 (Occurs Regularly) *N* = 246	47.81 (Occurs Occasionally) *N* = 196
2013–14[2]	*Individualized and Contingent Instruction****	51.47 (I Know This for Most of My Students) *N* = 212	48.23 (I Know This for Most of My Students) *N* = 177
	*Academic Learning Activities****	51.92 (Occurs Regularly) *N* = 212	47.80 (Occurs Occasionally) *N* = 185

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

[1] Numbers appear similar from one column to another but have different response categories because of the statistical technique used in analysis that assigns respondents a raw score on a scale of 1 to 100. Raw scores then are transformed into scale scores on the basis of the mean response for all survey participants that then are correlated to the response categories on the survey. The transformed scale may be very close to the cutoff point for a particular response category, so scores appear close, but response categories differ.

[2] Because staff members surveyed and interviewed each year are different, it is important not to compare 2012–13 data with 2013–14 data.

SOURCE: 21st CCLC Evaluation Activity Leader Surveys, 2012–13 and 2013–14.

In both 2012–13 and 2013–14, for items on the *Individualized and Contingent Instruction* construct, SPP and ACE-only activity leaders were similar in that the items most often selected were related to awareness of student strengths, recognition of the areas in which a student needed to improve, and a general understanding of how a

student feels about school. The least selected items were also similar for SPP and ACE-only activity leaders. These included knowing how a student performed on district or building-level assessments, awareness of the specific academic concepts and skills a student is struggling with, and knowing whether a student is struggling with academic content during the school day. This finding indicates that staff have more anecdotal knowledge of students' emotional needs than hard data on assessment test scores and performance.

Also in 2013–14, on the *Academic Learning Activities* items, the SPP activity leaders, more than the ACE-only activity leaders, reported directing students in activities associated with academic learning—for example, work with other students in small groups (in 2013–14, 27 percent versus 18 percent, respectively, said this *always occurs*), the use of computer-based learning programs (30 percent versus 14 percent, respectively, said this *always occurs*), review and practice of concepts taught during the school day (30 percent versus 18 percent, respectively, said this *always occurs*). See Appendix C, Table C-4 for individual response data.

The findings from this construct are likely because SPP programs were more academically focused than were ACE-only programs—the activity leaders, therefore, reported that they engaged students in *Academic Learning Activities* more often. The findings also suggest that both types of instructors were getting more anecdotal information about their students from the students themselves and perhaps their teachers than formalized information from assessments and grade reports. This corresponds to the finding from the *Use of Data* construct noted earlier in which both types of staff reported only occasionally accessing student data. This finding may suggest that both types of programs were not as connected to the schools that the students attend as they could be.

Youth Development Activities **Construct**

The activity leader survey included items related to *Youth Development Activities*. *Youth Development Activities* are those that give youth the opportunity to interact with other students to accomplish a shared purpose, to exercise choice in relation to their work and activities, and to engage in reflective activities in order to monitor their own learning. Research has shown that the presence of these activities is associated with a variety of positive social, emotional, and academic outcomes (Larson & Hansen, 2005; National Research Council, 2002). The Critical Success Model for ACE programs described in Chapter 1 and Appendix A include factors related to youth development (such as youth actively participating in their own learning and taking on leadership roles) and emphasize that such opportunities increase student engagement in afterschool activities.

In both 2012–13 and 2013–14, no statistically significant differences were found between ACE-only and SPP activity leaders on the provision of *Youth Development Activities* (see Table 7). In both cases, all types of activity leaders indicated providing *Youth Development Activities* regularly. This finding indicates that both types of activity leaders were implementing *Youth Development Activities* and that this is not a point of difference between the two types of staff.

Table 7. Mean Scale Scores on the *Youth Development Activities* Construct (2012–13 and 2013–14)

	Construct	SPP Mean[1]	ACE Mean[1]
2012–13	*Youth Development Activities*	49.88 (Occurs Regularly) *N* = 246	50.14 (Occurs Regularly) *N* = 194
2013–14[2]	*Youth Development Activities*	50.21 (Occurs Regularly) *N* = 212	49.76 (Occurs Regularly) *N* = 185

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$ - in this case there were no statistically significant findings.

[1] Numbers appear similar from one column to another but have different response categories because of the statistical technique used in analysis that assigns respondents a raw score on a scale of 1 to 100. Raw scores then are transformed into scale scores on the basis of the mean response for all survey participants that then are correlated to the response categories on the survey. The transformed scale may be very close to the cutoff point for a particular response category, so scores appear close, but response categories differ.

[2] Because staff members surveyed and interviewed each year are different, it is important not to compare 2012–13 data with 2013–14 data.

SOURCE: 21st CCLC Evaluation Activity Leader Surveys, 2012–13 and 2013–14.

Program Challenges

The activity leader survey also explored the types and frequencies of challenges that SPP and ACE activity leaders faced in delivering services to youth. It is useful to understand the challenges that different types of activity leaders faced for two reasons: (1) to contribute toward program improvement planning and (2) to understand whether the two types of programs were experiencing different challenges that may be reflective of the program. Nine types of challenges were included on the survey (see Appendix E, Table E-5 for a list of items).

In general, data show a lack of major challenges in both types of programs. However, there were some differences between SPP and ACE-only programming that are worth noting. As shown in Table 8, SPP activity leaders in both years, on average, had slightly higher scores on the various *Program Challenges* than did ACE-only activity leaders. In 2013–14, the difference (3.3 points on the scale) was slightly more pronounced than in 2012–13 (2.1 points). The difference was statistically significant, which means we can say with some degree of confidence that SPP activity leaders faced slightly more challenges than did ACE-only activity leaders in both years. However, in 2012–13 the responses indicate that challenges only *occurred occasionally*, and in 2013–14 the problems were reported as never occurring by ACE-only activity leaders and occurring occasionally by SPP activity leaders.

Table 8. Mean Scale Scores on the *Program Challenges* Construct (2012–13 and 2013–14)

	Construct	SPP Mean[1]	ACE Mean[1]
2012–13	*Program Challenges**	50.93 (Occurs Occasionally) *N* = 244	48.83 (Occurs Occasionally) *N* = 195
2013–14[2]	*Program Challenges****	51.53 (Occurs Occasionally) *N* = 211	48.25 (Never Occurs) *N* = 185

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

[1] Numbers appear similar from one column to another but have different response categories because of the statistical technique used in analysis that assigns respondents a raw score on a scale of 1 to 100. Raw scores then are transformed into scale scores on the basis of the mean response for all survey participants that then are correlated to the response categories on the survey. The transformed scale may be very close to the cutoff point for a particular response category, so scores appear close, but response categories differ.

[2] Because staff members surveyed and interviewed each year are different, it is important not to compare 2012–13 data with 2013–14 data.

SOURCE: 21st CCLC Evaluation Activity Leader Surveys, 2012–13 and 2013–14.

To understand the data above better, it is helpful to look at the individual challenges staff members faced. During 2013–14, the most common challenge that SPP activity leaders identified as *always occurring* was *Inconsistent Program Attendance* (4.7 percent), followed by *Instructional Time Interrupted by Student Behavioral Issues* (3.3 percent), and *Low Levels of Student Interest in Program Activities* (2.8 percent).

ACE-only activity leaders also identified *Instructional Time Interrupted by Student Behavioral Issues* (2.1 percent) and *Inconsistent Program Attendance* (1.7 percent) as the most common challenges. However, the third common challenge cited by ACE-only staff differed. In 2012–13, it was *Lack of Alignment Between School and OST Curriculum* and in 2013–14 it was *Difficulty Engaging Students in Curricular Materials or Program Content* (1 percent). One problem noted by SPP activity leaders—*Problems Getting Technology to Work Properly* (7 percent *always occurs* or *regularly occurs*)—was not reported as a problem by ACE activity leaders, who reported that this challenge *occasionally occurs* or *never occurs* (97 percent). See Appendix C, Table C-7 for individual item response data for 2013–14.

The results from this construct suggest that activity leaders are facing challenges only occasionally, a positive finding. The difference in the types of challenges reported makes sense given the different programming. Although youth behavior issues tend to be universal in OST programs, the other challenges are specific to the type of programming being offered. SPP programs often used computer-based interventions, meaning they are more regularly using technology than are ACE-only programs. Likewise, ACE-only programs are not focused specifically on academic support, so they may not have as many reasons to align their curriculum with that of the schools' students attend or to engage them in curricular materials.

Conclusions and Key Findings

This chapter focuses on addressing the following three RQs by examining quality at the organizational level for SPP and ACE-only programming implemented at ACE centers with SPP funding:

- **RQ1:** How does SPP programming compare with ACE-only programming in centers administered by the same grantee?
- **RQ3:** How does the quality of delivery differ between SPP and ACE-only programming?
- **RQ4:** What instructional and administrative practices lead to high student engagement?

Quality at this level is an important component of the SPP theory of change. As noted previously, the first two columns in the theory of change—Core Implementation Factors and Critical Student Experiences in SPP—happen at both the organizational level and the instructor or activity level. This chapter focuses on the former, attempting to understand how activity leaders use administrative practices such as sharing data, planning their activities, and addressing challenges to improve the quality of the student experience and how those practices differ across program types. Analyses of the staff survey data revealed some notable differences between the two types of programs but also highlighted that both SPP and ACE-only staff were engaging in high-quality practices—just in different ways. The differences between the two types of staff appear to reflect the orientation of the programs (with the SPP programs having more of a focus on academic intervention) and their associated structures.

Notable differences between SPP and ACE-only activity leaders include the following:

- **SPP activities appear to have more *Internal Communication* among activity leaders.** SPP activity leaders met more frequently (one meeting per week compared with monthly) than did ACE-only activity leaders. This difference may be due to staffing because the SPP programs tended to hire certified teachers from the feeder school to lead activities (as noted in Chapter 2), whereas the ACE-only programs hire youth development workers as well as certified teachers and community members. Teachers from the same school may have more opportunities to meet with one another because they are at the same location during the school day and they lead activities that are held at the school where they work. This is not the case with activity leaders who are not from the school, who are housed at the grantee organization, or who hold another position. In addition, it may be that SPP activity leaders were more focused on implementing common academic interventions as well as getting new programming up and running so are more in need of common planning time than are ACE-only activity leaders who are implementing more disparate types of programming.
- **SPP activity leaders were more likely than the ACE-only activity leaders to plan activities that address the development of specific academic skills.** In particular, SPP activity leaders tended to plan activities that explicitly promote

skill building in relation to state standards. This difference between the SPP and ACE-only activity leaders reflected the purpose of the SPP programs, which was to assist at-risk students in developing skills that are measured on the STAAR.

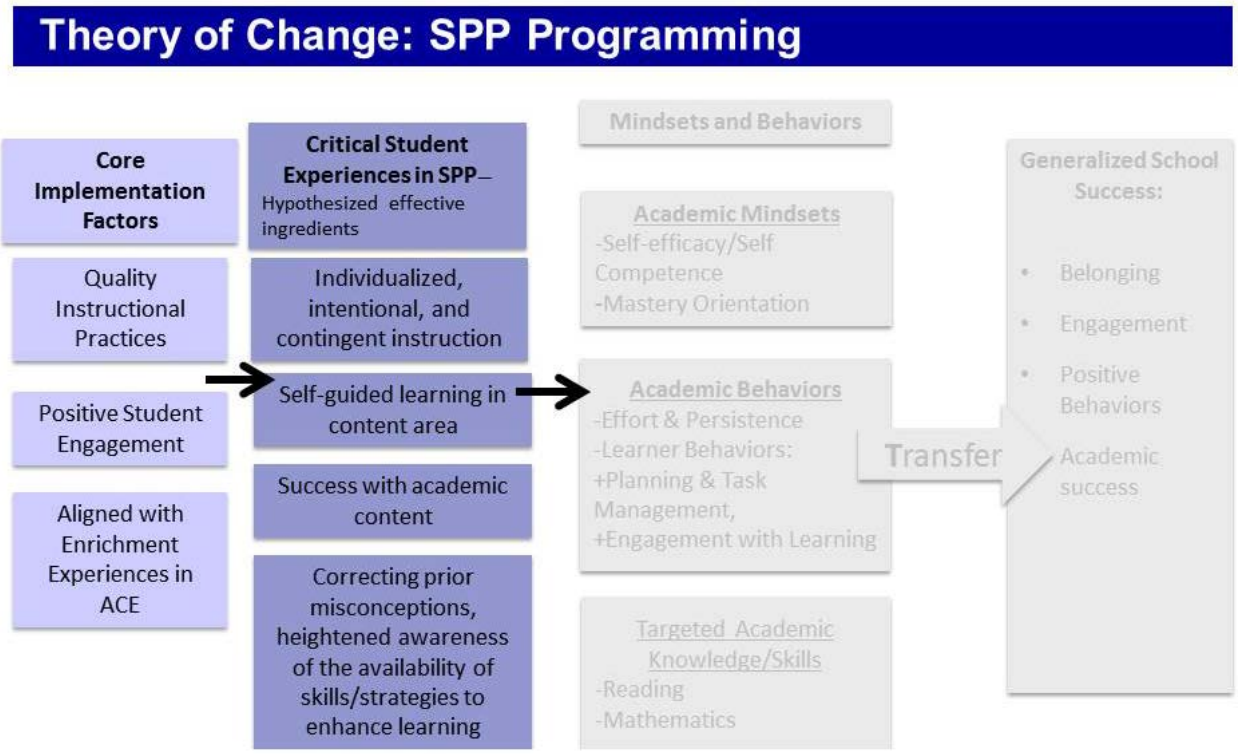
- **SPP activity leaders directed students in *Academic Learning Activities* more frequently than did ACE-only activity leaders.** This difference again reflects the different purposes of the programs, one program being more focused on explicit academic instruction than the other program.
- **Although challenges were minimal, SPP programs faced more challenges than did ACE-only programs.** The challenges reported by SPP activity leaders were different and more numerous than challenges reported by ACE-only activity leaders. This may be partly because the SPP programs were newer and, therefore, may have had implementation challenges. The nature of the differences—SPP activity leaders faced technology and attendance challenges, and ACE-only activity leaders faced school alignment challenges—may reflect the nature of each of the program types (e.g., SPP programs were more likely to use technology than were ACE-only programs, so they might face more technological challenges than did ACE-only programs).

Overall, the findings from analysis of staff surveys indicate that SPP and ACE-only activity leaders are both engaging in high-quality organizational practices that, according to the SPP theory of change, support critical student experiences. Although this chapter does not yet connect those administrative practices to student engagement, analysis of the survey results reveals that activity leaders engaged in planning and communication practices that may contribute to student engagement. In addition, activity leaders from each type of programming differed in their approaches to instruction (e.g., SPP activity leaders focused more on planning academic activities), but primarily the activity leaders differed in ways that reflect their program types.

Chapter 4: Quality and Engagement at the POS

Throughout the past three years, the issue of program quality has been a consistent theme of the statewide evaluation of the ACE programs in Texas. One of the expectations the evaluation has tested has been that higher quality programming would result in higher student participation rates in the ACE programs and higher student engagement in the program activities. This expectation is evident in the SPP theory of change, which posits that high levels of student engagement in activities will contribute to students' improved academic outcomes. Therefore, it is important to examine student engagement to determine how it varies across activities and to identify the features of the more engaging activities. This chapter examines the first two columns of the SPP theory of change, with a particular focus on connecting quality instructional practices and student engagement to critical student experiences in SPP.

Figure 5. SPP Theory of Change, Chapter 4



Chapter 4 explores program quality at the POS and begins to test this theory that program quality relates to student engagement by addressing the following three RQs:

- **RQ3:** How does the quality of delivery differ between SPP and ACE-only programming?
- **RQ4:** What instructional and administrative practices lead to high student engagement?

- **RQ5:** How engaged are young people in SPP and ACE-only programming? What is the relationship between quality of delivery and student engagement?

Chapter 3 began to answer RQ3 and RQ4 by looking at administrative practices in SPP and ACE-only programming and which administrative practices appeared to be related to student engagement. In this chapter, the evaluation team takes those two sets of analyses to the next level, looking at how instructional practices varied in the two types of programming and the connection between quality and engagement.

In order to answer the RQs as well as begin to test the theory of change, the evaluation team conducted two different sets of analyses related to program quality and engagement, one analysis in 2012–13 and one analysis in 2013–14.

In 2012–13, the evaluation team developed a program typology for SPP programs focused on two dimensions of programming: the mode of delivery (how content is delivered) and the program approach to improving skills. These descriptions, based on data collected during site visits, then were used to ground the subsequent discussion of student engagement and the type of activities in which students are engaged most actively to examine better the accuracy of the SPP theory of change.

In 2013–14, the evaluation team took the analysis connecting program quality and engagement to a deeper level, examining the connection between quality and youth engagement by using a series of sophisticated statistical techniques to conduct analyses connecting student engagement surveys and observational data from site visits on the quality of programs. These site visits included both ACE centers with SPP funding and a sample of similar ACE-only centers within the same grantee that did not have SPP funding. Observation data as well as student engagement data from these 26 sites were used in the analyses described in the following section to understand better the difference between SPP and ACE-only programs.

SPP Typology Analysis 2012–13

The SPP typologies were developed based on data collected from observing 22 activities during one-day site visits, in spring 2013, to 15 centers representing each of the 15 grantees that received SPP funding. The typologies were based on program activities—how they were designed—rather than on other features, such as the grade levels of students served or the subject area. As noted in the theory of change, it is essential that activities are designed with core implementation factors in mind in order to provide students with a set of critical experiences. Thus, the SPP typologies were created based on the types of experiences SPP program activities offered students. Only SPP programs were included in the typology analysis because the programming was new, and our goal for the evaluation was to understand better the instructional practices being employed in those program activities. The research team identified two categories of typologies that indicate how students might experience the activities (both with respect to what they do and what they learn). The first typology is associated with mode of delivery—the general structure of the sessions that directs what students do.

The second typology is associated with the overall curricular approach of the activities—what students learn.

Mode of Delivery Typology

Two modes of delivery were identified: computer-based delivery and face-to-face delivery. Computer-based modes of delivery were those in which computer-based curricula for reading (e.g., Istation, Achieve3000, Power Reading, Reading Plus) or mathematics (e.g., Think Through Math, Essential Skills) were used to assess and improve students' subject-area skills. Face-to-face modes of delivery were those in which an activity leader worked with small groups of students on academic skills and learning strategies and students did not use a computer-based program. The majority of SPP programs used a combination of computer-based and face-to-face delivery modes. In summary, the SPP programs fell into three categories as follows:

- **Combined Mode (Computer-Based Delivery and Face-to-Face Delivery).** Nine (60 percent) of the 15 SPP centers visited used both types of formats within the same activity. In these SPP centers, either some activities were computer-based and other activities were face-to-face or an instructor actively facilitated computer-based sessions by directing student activities and providing consistent coaching and feedback.
- **Face-to-Face Delivery Only.** Three (20 percent) of the SPP centers used only face-to-face interventions and did not use computer-based learning programs.
- **Computer-Based Delivery Only.** Three (20 percent) of the SPP programs provided only interventions through computer-based programs, with minimal or no activity leader facilitation.

Among the 22 activities that were observed at the 15 centers with SPP programs, 11 activities used a face-to-face mode of delivery, seven activities used an online computer-based mode of delivery, and four activities used a combination of computer-based and face-to-face modes of delivery.

Approach Typology

In addition to the mode of delivery for the activities, the SPP programs demonstrated two different approaches for improving students' academic performance. Those approaches were as follows:

- A ***Learning Strategies*** **approach**, in which specific curriculum guided the application of learning strategies that potentially could be applied to improve study skills broadly across various subject areas. Five of the 15 SPP programs observed in spring 2013 were identified as using a *Learning Strategies* approach. Four of the five based programming on a curriculum developed by a university or other research organization.[18]

[18] The science, technology, engineering, and mathematics–based curriculum for the SPP program in Harlingen Consolidated Independent School District was developed by the district.

- A ***Skills-Building*** **approach**, in which a curriculum, usually a computer-based program, was used to develop specific subject-area skills, and/or small-group instruction focused on developing specific skills. The programs usually had an assessment component through which both students and activity leaders could track student progress. Ten of the 15 SPP programs observed in spring 2013 were identified as using a *Skills-Building* approach.

Learning Strategies **Approach.** The *Learning Strategies* approach, although used in activities specifically focused on one subject area (e.g., reading), emphasized teaching students a process for learning, which students could apply to different subject areas and learning experiences. For example, one of the curricula focused on the importance of metacognition, in which students think about their thinking and how they solve problems. Another curriculum emphasized connecting learning to real-world activities. A third curriculum was designed to equip students with various learning strategies to help them understand information and solve problems. A fourth curriculum had students engage in a five-step learning process. These examples are described in greater detail in Appendix F.

Skills-Building **Approach.** The majority (67 percent) of the SPP programs emphasized a *Skills-Building* approach in their academic sessions. This is not surprising given the direction SPP sites received to implement curricula designed to improve specific mathematics and reading skills among student participants. A combination of computer-based and face-to-face sessions were implemented for eight of the *Skills-Building* SPP programs where site visits were conducted. Three SPP programs, all serving secondary students, offered computer-based sessions exclusively at the time of the site visits. One SPP program that served elementary students used face-to-face sessions exclusively.

Typologies Summary

As the formats and approaches show, SPP programs offered students different types of learning experiences within the program activities. The programs differed in terms of format and how they combined face-to-face and computer-based learning experiences as well as the general approach—whether the curricula focused mainly on learning strategies or content-area skills building.

This analysis leads to several questions: Are students more engaged in activities of one type more than other types? For example, are students more engaged in a computer-based format or a face-to-face format? Are students more engaged in activities with a *Learning Strategies* approach than in activities with a *Skills-Building* approach? To what extent do instructional practices influence student engagement? The evaluation team used an observation protocol to examine whether specific instructional practices—the creation of a supportive environment and positive interactions—were present in a series of program observations. The discussion related to quality of activities (which follows in the next section) focuses on student engagement and examines the general characteristics of the activities in which students were more or less engaged, including type of activity.

Exploring the Linkage Between Program Quality and Engagement

Background and Context 2010–11 Through 2012–13

The instructional practices adopted by OST staff can have a substantial influence on youth experience in programs. In any given individual activity session, the particular instructional practices staff select either will be conducive to creating and sustaining interest and engagement in the activities that are being provided or will not sustain youth interest and engagement will be low. The literature suggests that activities that help young people feel challenged, make content relevant to participating youth, make youth feel like they are learning something, and generate positive effects are more likely to support participant interest and engagement (Assor, Kaplan, & Roth, 2002; Csikszentmihalyi, 1990; Csikszentmihalyi & Schneider, 2000; Hidi & Renninger, 2006; Renninger & Hidi, 2011). Engagement is essential to contributing to youth outcomes. Research has shown that youth who are highly engaged in programming are more likely to experience improvement in a variety of outcomes (Kataoka & Vandell, 2013). It is this literature base that formed the foundation for the SPP theory of change that hypothesizes a connection between high-quality practices; youth engagement; and, ultimately, youth outcomes.

At this point, it is helpful to refer back to the 2010–11 and 2011–12 evaluation reports to set the context for the 2012–13 and 2013–14 analyses (Naftzger et al., 2012; Naftzger et al., 2013). In the first two years of the evaluation, steps were taken to measure the presence of instructional practices in ACE activities[19] likely to support youth interest, engagement, and skill development through direct observation of programming by using quality assessment tools commonly used in the OST field. These tools included the Program Quality Assessment (PQA),[20] which focuses on *process* quality (e.g., the extent to which staff provide supports and opportunities that result in the creation of a supportive, interactive, and engaging environment) and the *Academic Skill Building* section of the Assessment of Program Practices Tool–Observation (APT-O), which outlines the types of support and opportunities OST staff can provide to support skill development in particular *content* areas (see the Observation Measures section in Appendix B for more detail on these two instruments). In addition, the evaluation team attempted to measure youth engagement in the program by using two approaches: (1) an end-of-session youth survey administered to students in Grades 4 and above at the conclusion of an observation asking students to report on their levels of interest and engagement during the observed session (see Appendix E, Table E-6 for the survey items) and (2) use of an observation-based behavioral measure of youth engagement, the Observation of Child Engagement (OCES) (see the Observation Measures section in Appendix B for more information on the OCES). The evaluation team hypothesized that greater adoption of the instructional practices outlined in the PQA and APT-O would be associated positively with both youth engagement in programming as measured by

[19] At this point in the evaluation, SPP did not exist, and only ACE activities were included in the observations.

[20] Two versions of the PQA were used. For programs serving students in Grades 6–12, the Youth PQA was used. For those serving students in Grades K–5, the School Age PQA was used.

the youth survey and the OCES and the cultivation of the types of youth outcomes desired from participation in OST offerings.

Through use of these tools, the evaluation team examined the relationship between program quality as defined by the criteria outlined in the PQA and engagement. The practices assessed in the PQA support the creation of a supportive, interactive, and engaging environment that helps foster youth feeling a sense of agency, efficacy, belonging, and encouragement. In 2010–11, a significant and positive relationship was found between the total PQA score and youth-reported engagement on the end-of-session student survey. This finding suggested that the more staff adopted practices and created opportunities for students on the basis of the criteria outlined in the PQA, the higher the engagement students reported on the survey. In 2011–12, the evaluation team explored the relationship between program quality in ACE-funded activities and school-related outcomes. The goal of these analyses was to answer the following question: Does the impact on student outcomes vary by relevant ACE program characteristics, including center quality? Centers that (a) were observed providing extensive academic content and supportive, interactive, and engaging activities and (b) were characterized by higher levels of observed youth engagement demonstrated more of an impact on reducing disciplinary incidents and on increasing student grade promotion.[21] In addition, centers demonstrating less of a capacity to provide activities with these characteristics demonstrated less of an impact on supporting student performance on the TAKS-Reading/ELA assessment.

Results from both early years of the evaluation support the theory of change for OST programs outlined in Figure 1 in Chapter 1 and the more specific theory of change developed for SPP programs outlined at the beginning of this chapter. Although the analyses in those years did not assess fully the relationship between each of the steps articulated in the model, there is an apparent connection between higher quality programs, increased student engagement, and improved youth outcomes in ACE programs overall.

As outlined previously, in 2012–13, the evaluation team observed SPP activities and classified the activities both by mode of delivery and by general approach. In addition, at the completion of the observed activity sessions, the PQA was scored by the observers to obtain an estimate of program quality for the observed activities (see Appendix G for a table of quality scores).[22] Some differences in PQA scores were found, depending on the format and approach used to deliver activities, including the following:

[21] In conducting these analyses, the evaluation team compared ACE program participants with similar youth not participating in the program but attending the same schools as the ACE participants. Program quality variables were based on PQA, APT-O, and OCES scores obtained from the activities observed by the evaluation team. The team used correlational analyses to assess the relationship between a center's quality status and each of the center-level effect size estimates. In this sense, although a relationship can be said to exist between program quality and youth outcomes, it cannot be said that the level of program quality caused these outcomes to happen.

[22] At the time of observation, the program activities had been implemented only recently, and, in numerous SPP centers, activity leaders had completed only recently, or were still engaged in, professional training related to the programming. In some of the SPP programs, instructional materials and computer-based programs had become available only recently. Therefore, the quality scores should be interpreted with caution.

- All of the activities with a combined format (computer based and face-to-face) had relatively high instructional PQA scores[23] (from 53 to 72 on a 100-point scale) and were in the top 40 percent of the activities observed. Four of the six activities with the highest PQA *Engagement* ratings were delivered in a combined format. The *Engagement* section of the observation tool represents the most difficult practices to implement, indicating that these four programs were of very high quality.
- Among the 11 activities with the highest PQA engagement scores, all were delivered in a face-to-face or combined format. In contrast, all seven of the observed activities that were delivered in only a computer-based format were among the 11 activities with lowest PQA engagement scores, indicating that the computer-based format is less conducive to student engagement.
- The three activities with the highest instructional PQA scores (65, 72, and 73, respectively) used a *Learning Strategies* approach, for which there was an established curriculum. These three activities were all in elementary schools. However, the activity scored lowest on engagement also used a *Learning Strategies* approach, although the activity did not follow a specific curriculum and was offered to high school students. This finding could indicate that elementary programs tend to be higher quality (all of the high school programs observed had average scores below 40), or it could indicate that use of an established curriculum is a higher quality approach. More observations would be needed to determine which explanation is most accurate.

In addition to the PQA scores, the youth engagement survey yielded information about the quality of activities with different modes and approaches. Examples of questions appearing on the survey included *Was today's activity interesting?* and *Did you enjoy what you were doing during this activity?* In responding to questions on the youth engagement survey, students endorsed one of three response options: (1) *yes, very much*, (2) *sort of*, and (3) *not at all*. Youth-reported engagement was higher among students in the combined-format activities than among students in single-format activities. More specifically, 28 percent of survey respondents in the combined format had a survey score that put them in the *yes, very much* response category. Having a survey score in this portion of the rating scale was indicative of a high level of engagement in the activity session being observed. In comparison, only 10 percent of survey respondents in the computer-based group had a survey score that put them in this category, and 17 percent of youth in the face-to-face categories had a survey score that put them in this high engagement category. Results from the youth engagement survey indicated that students were most engaged in activities that used a combined format. These results also were consistent with the PQA scores for such activities, which were higher than the other activity formats.

[23] Instructional PQA score is an average score for the three domains of the PQA related to instruction: *Supportive Environment*, *Interaction*, and *Engagement*.

The findings from the first three years of the evaluation informed the data collection activities and data analysis for 2013–14. A goal of the 2013–14 evaluation was to assess each of the pathways specified in the theory of change, starting with program quality and ending with a series of desirable school-related outcomes among youth participating intensively in programming (this latter will be examined in 2014–15). In 2014–15, during the fifth year of this evaluation, the research team will continue to test the SPP theory of change by exploring the relationship between levels of program engagement and the cultivation of youth skills and beliefs and transfer outcomes related to academic achievement and behaviors. This analysis will build on the quality/engagement analysis conducted in 2013–14, described in the following sections.

Connecting Quality and Youth Engagement: 2013–14[24]

As noted earlier, during the spring semester of 2014, members of the evaluation team conducted two-day site visits at 26 sites—13 ACE centers that had SPP funding and 13 ACE-only centers. During these visits, members of the evaluation team observed a total of 89 afterschool activities serving youth in Grades K–12. In general, there were not significant differences in the observed level of program quality between SPP and ACE-only activities, and programs tended to score at about 54 out of 100 on the PQA (see Appendix D, table D-2 for more detail). The activities observed were primarily either academic enrichment (73 percent) characterized by an intentional effort to build youth skills in a specific academic content area or nonacademic enrichment (26 percent), which were more apt to emphasize the provision of developmentally appropriate activities that supported youth development more broadly. The median number of youth participating in an observed activity was 10 (ranging from 1 to 60), and the median number of staff was 1 (ranging from 1 to 5). As shown in Tables 9a and 9b, activities observed in SPP centers were more likely to be classified as academic enrichment ($p < 0.05$, chi-square = 8.851, df = 2) and have more staff present on average ($p < 0.001$, $t = 4.498$, df = 55.646) than were activities observed in ACE-only centers.

[24] There are several limitations associated with the findings described in this section. First, the evaluation team only partially explored the domain of variables that may have had an impact on degree to which youth were engaged in observed activities. It is possible that other critical variables not included in the analyses summarized here are mediating the relationship between activity design and delivery and the level of youth engagement. In other words, although the findings described in this section of the report demonstrate a relationship between certain characteristics and practices associated with OST program delivery, we cannot say definitively that a given characteristic or practice caused either higher or lower levels of youth engagement.

Table 9a. Activity Type Associated With Observed Activities by Center Type (2013–14)

Activity Type	Activities in SPP Centers[1]		Activities in ACE Centers[2]	
	N	%	N	%
Academic enrichment	39	86.7%	26	59.1%
Nonacademic enrichment	6	13.3%	17	38.6%
Tutoring	0	0.0%	1	2.3%

[1] *N* = 45.
[2] *N* = 44.

SOURCE: 21st CCLC Evaluation Site Visit Observation Notes (2013–14).

Table 9b. Activity Participants Associated With Observed Activities by Center Type (2013–14)

Youth and Staff Participants	Activities in SPP Centers[1]	Activities in ACE Centers[2]
Mean number of youth	10.75	10.07
Mean number of staff	1.64	1.09

[1] *N* = 45.
[2] *N* = 44.

SOURCE: 21st CCLC Evaluation Site Visit Observation Notes (2013–14).

Exploring the Relationship Between Activity Characteristics and Quality

In order to explore the relationship between activity characteristics and quality and levels of youth engagement, we conducted a series of analyses by using engagement scores from both the OCES and the youth survey as outcomes. Each of the analyses used the following variables related to activity characteristics and quality as predictors:

- *SPP Center.* Indicated whether the center was funded by SPP. This variable was included to explore whether the level of engagement varied between SPP and ACE-only centers.
- *Academic Enrichment Activity.* Indicated whether the activity being observed was classified as academic enrichment. It was hypothesized that activities with more explicit academic content may be less engaging for participating youth.
- *Staff-to-Student Ratio.* Represented the ratio of staff to youth participating in an observed activity. Activities for which this ratio was lower were hypothesized to be associated with greater youth engagement.
- *High Reading Content.* Indicated whether the activity was (a) intentionally focused on building skills in reading and writing and (b) above the 75th percentile on the APT-O scale score for either reading or writing content. Again, it was anticipated that the presence of more academic content would be related negatively to youth engagement.

- *High Mathematics Content.* Indicated whether the activity was (a) intentionally focused on building skills in mathematics and (b) above the 75th percentile on the APT-O scale score for either mathematics problem solving or communication. See the prior bullet in terms of why this variable was included in the analyses related to engagement.
- *Total PQA Score.* The mean scale score associated with six subscales used to score the PQA. Steps taken to score the PQA were different from the approaches employed in preparing the interim and year two reports. Previously, either a total PQA score was calculated or domain-level scores were calibrated (e.g., separate scores for the *Supportive Environment*, *Interaction*, and *Engagement* domains). For this analysis, analysts created subscale scores to explore better whether specific types of PQA practices were found to be related especially to youth engagement. In undertaking analyses to create subscale scores, we found that some subscales demonstrated poor reliability, given a lack of variation in scores across activities (almost all activities scored highly on these), and we therefore dropped them from further analyses. Other subscales with poor reliability but good variability in scores were combined with other subscales to construct a usable scale. Ultimately, six subscale scores were calibrated from PQA data composed of the following sections of the instrument:

  ACTIVITY CHARACTERISTICS THAT SUPPORT ACTIVE ENGAGEMENT

  - Activities support active engagement (*Supportive Environment* subscale H).

  ENCOURAGEMENT and FEEDBACK

  - Staff support youth in building new skills (*Supportive Environment* subscale I).
  - Staff support youth with encouragement (*Supportive Environment* subscale J).

  COOPERATIVE LEARNING and BELONGING

  - Youth have opportunities to develop a sense of belonging (*Interaction* subscale L).
  - Youth have opportunities to collaborate (*Interaction* subscale M).
  - Youth have opportunities to act as group facilitators and mentors (*Interaction* subscale N).

  POSITIVE ADULT INTERACTIONS

  - Youth have opportunities to partner with adults (*Interaction* subscale O).

  PLANNING and CHOICE

  - Youth have opportunities to set goals and make plans (*Engagement* subscale P).
  - Youth have opportunities to make choices based on their interests (*Engagement* subscale Q).

REFLECTION

- Youth have opportunities to reflect (*Engagement* subscale R).

- *Youth Grade Level.* Analyses in which the OCES was the outcome of interest included a variable that indicated whether the observed activity was being attended by youth in Grades K–3. In cases in which the student survey engagement score was the outcome of interest, the grade level of the student survey respondent was included in the analysis. It was hypothesized that the relationship between program quality and youth outcomes may vary by the grade level of youth served.

Observed Engagement Scores. First the team tried to understand whether activity characteristics (see Appendix D, Table D-4 for a description of characteristics) and quality might be related to the level of youth engagement observed by site visitors and measured by the OCES. In order to do this, the evaluation team ran a series of statistical tests (multiple regression analyses) to explore how activity characteristics and quality were related to the level of observed youth engagement in a given activity. As shown in Table 10, lower staff-to-student ratios were associated positively with the levels of observed youth engagement in activities, as were higher total PQA scores, which indicated that staff leading these activities were adopting practices associated with the creation of supportive, interactive, and engaging learning environments. Each of these predictors had a statistically significant relationship with the level of observed student engagement.

Table 10 shows that most of the other characteristics included in the analysis were related negatively to OCES scores. What this means is that, in most cases, the characteristic was associated with lower levels of observed youth engagement on the basis of the PQA score—that is, activities that were designated as SPP, that were designated as academic enrichment, that had high mathematics or reading content, and that were at the K–3 level all had lower levels of observed youth engagement. These differences were not significant, which means that we cannot determine with a high level of confidence that the characteristic predicted the lower OCES scores. There may be some other factor that contributed to the lower OCES scores.

Table 10. Estimated Effect of Various Activity Characteristics and Quality Measures on Observed Engagement Scores: Activities Serving Youth in Grades K–12 (2013–14)

Predictor	Unstandardized Coefficient	Standard Error
SPP Center	–4.304	3.983
Academic Enrichment Activity	–1.208	5.098
Staff-to-Student Ratio	31.094**	12.226
High Reading Content	–3.965	4.626
High Mathematics Content	–1.227	4.962
Total PQA Score	0.402**	0.193
Youth Grade Level K–3	–1.011	3.873
*p < 0.10, **p < 0.05, ***p < 0.01.		

SOURCE: 21st CCLC Evaluation PQA and APT-O observation scores for 89 observed activities, 2013–14 and TEA, TX21st, 2013–14.

The OCES was used to observe engagement during all activities, regardless of student grade level. The first set of analyses described in Table 10 was for all activities and all grades. Activities serving Grades 4–12 had both the OCES and the self-reported youth engagement surveys as sources of information on engagement. Given that OCES scores were the only measure of engagement for activities serving youth in Grades K–3, the analysis outlined in Table 10 was rerun but only with activities that served youth in Grades K–3 to see whether there was a difference.

As shown in Table 11, the results were similar to those outlined in Table 10, with two differences. First, staff adoption of practices to support the creation of supportive, interactive, and engaging learning environments as represented by the total PQA score was no longer significantly related to the level of observed youth engagement. This finding was not unexpected, given that many of the practices outlined on the PQA focus on youth development practices that appear easier to implement with older elementary, middle, and high school youth. Second, SPP activities were related significantly and negatively to the observed engagement of youth. This finding was consistent with what was hypothesized, given that SPP activities were focused more explicitly on academic content than were ACE-only activities. Lower staff-to-student ratios still were associated significantly and positively with the levels of observed youth engagement in activities. The strength of this relationship was even greater for activities serving youth in Grades K–3 than what was demonstrated when all observed activities in Grades K–12 were considered in Table 10. This set of analyses indicates that a low staff-to-student ratio is even more important for the youngest students than for older students; that SPP activities were less engaging to the youngest students; and that the overall program quality based on the adoption of practices that foster youth feeling a sense of agency, efficacy, belonging, and encouragement may not have been as connected to youth engagement for students in Grades K–3 as for older students.

Table 11. Estimated Effect of Various Activity Characteristics and Quality Measures on OCES Scores: Activities Serving Youth in Grades K–3 Only (2013–14)

Predictor	Unstandardized Coefficient	Standard Error
SPP Center	–10.444**	4.937
Academic Enrichment Activity	–1.980	6.448
Staff-to-Student Ratio	135.451***	35.429
High Reading Content	–6.792	5.593
High Mathematics Content	–3.257	5.750
Total PQA Score	0.152	0.294
*p < 0.10, **p < 0.05, ***p < 0.01.		

SOURCE: 21st CCLC Evaluation OCES observation scores, 2013–14; and TEA, TX21st, 2013–14.

Taken together, the analyses of data on the observed level of youth engagement in programming tell us that youth appear to be more engaged in programs with a low staff-to-youth ratio and high levels of program quality.

Student Self-Reported Engagement Survey Scores

The OCES scores and their connection to activity characteristics tell one part of the story about youth engagement. A second part of the story is youth-reported engagement collected via student engagement surveys that were distributed at the end of program observations during the spring 2014 site visits. A total of 401 student engagement surveys were collected during observations of 49 SPP and ACE-only activities[25]. The observation team used a different statistical technique to explore the relationships among activity-level characteristics (see Appendix D, Table D-4), instructional quality, and the level of reported youth engagement in the activity session observed. Given the nested nature of these data (youth nested within activities and activities nested within centers), a series of hierarchical linear models were run to explore the relationship between center- and activity-level characteristics and quality and the level of reported youth engagement in the activity session observed. By using hierarchical linear modeling (HLM) to conduct these analyses, it was feasible to assess how variables at the center- and activity-levels may be related to self-reported youth engagement in activities (see Appendix D for more detail on the method). The results of the analyses to examine the connections among youth engagement survey results, activity characteristics, and overall quality scores are summarized in Tables 12–15.

The first model that was run explored the relationship between center- and activity-level characteristics and youth-reported engagement in activities. Only one center-level

[25] Only 49 activities are included in this sample because youth surveys were not distributed at all activities. Only youth at activities serving Grades 4–12 received the survey. Also, in some centers, youth surveys were not possible directly after program observation because of timing or staffing conflicts.

characteristic was included in the model—the SPP status of the center. Activity-level characteristics included in the model were whether or not the activity was classified as academic enrichment (as opposed to nonacademic enrichment) and the staff-to-student ratio. The only student-level predictor included in the model was student grade level. In this initial model, only the main effects of the center, activity, and youth-level characteristics included in the model were assessed. The *main effect* is the effect of an independent variable (e.g., SPP status of the center, staff-to-student ratio) on the level of youth engagement, averaging across the levels of the other independent variables included in the model.

As shown in Table 12, two of the characteristics included in the model were related significantly and negatively to engagement. Youth reported lower levels of engagement when participating in academic enrichment activities (compared with participation in nonacademic enrichment activities). This finding is consistent with what was hypothesized, given an expectation that youth would demonstrate lower levels of engagement when the provision of academic content within an activity session was more overt. In addition, youth were less engaged in activities observed in SPP centers (this was a moderately significant relationship). Here again, this finding may be related to the goals of the SPP program, which focused more intentionally on the identification and implementation of intervention strategies meant to support youth academic development and growth. This may have resulted in activities that were less enjoyable, interesting, and open-ended than activities provided in centers receiving only ACE funding.

Table 12. Estimated Main Effects of Various Activity Characteristics on Student Survey Engagement Scores (2013–14)

Predictor	Unstandardized Coefficient	Standard Error
Main Effects		
SPP Center	–4.426*	2.229
Academic Enrichment Activity	–7.756***	2.364
Staff-to-Student Ratio	10.020	13.033
Student Grade Level	–0.527	0.420
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.		

SOURCE: 21st CCLC Evaluation; PQA and APT-O observation scores from 49 activities; 401 youth engagement surveys; and TEA, Public Education Information Management System (PEIMS), 2013–14.

In addition to assessing the main effects of center- and activity-level characteristics on youth engagement, a second model was run using the same predictors to explore how the predictors may have interacted to influence engagement. For example, as shown in Table 13, a significant interaction was found between the SPP status of a center and the academic enrichment status of an activity. This finding means that the negative relationship between academic enrichment activities and levels of youth engagement were moderated by the SPP status of the center. In this case, the difference in the level

of youth engagement between academic enrichment and nonacademic enrichment activities was *smaller* in SPP centers compared with that in centers receiving only ACE funding where the difference between academic and nonacademic enrichment activities was larger. In this sense, the level of youth-reported engagement in academic and nonacademic enrichment activities was more consistent in SPP centers than in centers receiving only ACE funding.

Other significant interactions are outlined in Table 13 related to student grade level. Typically, youth engagement in 21st CCLC programming declines as youth get older. The significant interaction between SPP status and student grade level outlined in Table 13 demonstrates that the decline in engagement as youth grade increases was less intense in SPP centers than in centers funded only by ACE, except when the activity is an academic enrichment activity. In this case, the decline in youth engagement as grade level increases was more substantial in SPP centers than in ACE-only centers. These findings suggest there were some nonenrichment activities provided in SPP centers that were especially engaging as youth grade level increased.

Table 13. Estimated Interactions of Various Activity Characteristics With Student Survey Engagement Scores (2013–14)

Predictor	Unstandardized Coefficient	Standard Error
Cross-Level Interaction of SPP Center Status With Activity Characteristics		
SPP Status x Academic Enrichment Activity	–15.370*	9.066
SPP Status x Staff-to-Student Ratio	0.681	20.703
Cross-Level Interaction of Student Grade Level With Activity Characteristics		
Student Grade Level x SPP Status	–14.000*	6.908
Student Grade Level x Academic Enrichment Activity	–1.507	1.377
Student Grade Level x Staff-to-Student Ratio	1.334	7.684
Cross-Level Interaction of Student Grade Level With SPP Center Status and Activity Characteristics		
Student Grade Level x Academic Enrichment Activity x SPP Status	12.625*	7.060
Student Grade Level x Staff-to-Student Ratio x SPP Status	–3.501	10.356
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.		

SOURCE: 21st CCLC Evaluation; PQA and APT-O observation scores from 49 activities; 401 youth engagement surveys; and TEA, PEIMS, 2013–14.

In addition to center- and activity-level characteristics, the relationship between quality scores obtained from activity observations and youth engagement also was assessed.

In these models, activity-level variables were included that indicated (a) the extent to which activity leaders adopted practices to create a supportive, interactive, and engaging environment (the total PQA score) and (b) the presence of high reading and mathematics content (based on APT-O scores). Examination of how these activity-level variables were related to youth-reported engagement is important because these variables represent attributes of activity design and delivery that can be modified in the future to support the level of youth engagement in afterschool activities better. The SPP status of the center and student grade level also were included in these models.

The first model that was run to assess the relationship between measures of program quality and engagement focused on estimating the main effects of the predictors included in the model. As shown in Table 14, three of the characteristics included in the model were related significantly to engagement. Here again, levels of youth engagement were lower in activities provided in SPP centers. In addition, engagement was also lower in activities observed to have high levels of reading content, although this was a moderately significant relationship. This finding was consistent with the hypothesis that the presence of more overt academic content would depress self-reported levels of youth engagement. Finally, youth engagement was significantly higher in activities that also received higher total PQA scores, indicating that staff leading such activities were more apt to adopt practices that supported the creation of a supportive, interactive, and engaging environment. This finding was also consistent with what was hypothesized.

Table 14. Estimated Main Effects of Various Quality Practices on Student Survey Engagement Scores (2013–14)

Predictor	Unstandardized Coefficient	Standard Error
Main Effects		
SPP Center	–4.414**	2.102
High Reading Content	–4.597*	2.278
High Mathematics Content	–3.888	2.644
Total PQA Score	0.234**	0.088
Student Grade Level	0.206	0.574
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.		

SOURCE: 21st CCLC Evaluation; PQA and APT-O observation scores from 49 activities; 401 youth engagement surveys; and TEA, PEIMS, 2013–14.

A second model was run using the same predictors outlined in Table 14 to explore how variables related to program quality may have interacted to influence engagement. As shown in Table 15, only one interaction was significant. In this case, the SPP status of the center was found to moderate the relationship between the total PQA score and youth-reported engagement. More specifically, in SPP centers, the positive relationship between higher PQA scores and self-reported youth engagement was less intense than in centers funded only by ACE. In this sense, greater adoption of practices designed to

create supportive, interactive, and engaging learning environments was less likely to be related to engagement in activities provided in SPP centers.

Table 15. Estimated Interaction of Various Quality Practices With Student Survey Engagement Scores (2013–14)

Predictor	Unstandardized Coefficient	Standard Error
Cross-Level Interaction of SPP Center Status With Quality Practices		
SPP Status x High Reading Content	–10.531	8.589
SPP Status x High Mathematics Content	–3.184	5.970
SPP Status x Total PQA Score	–0.445**	0.186
Cross-Level Interaction of Student Grade Level With Quality Practices		
Student Grade Level x SPP Status	–1.007	1.725
Student Grade Level x High Reading Content	6.682	4.910
Student Grade Level x High Mathematics Content	2.529	2.732
Student Grade Level x Total PQA Score	–0.065	0.054
Cross-Level Interaction of Student Grade Level With SPP Center Status and Activity Characteristics		
Student Grade Level x High Reading Content x SPP Status	–6.882	5.170
Student Grade Level x High Mathematics Content x SPP Status	–4.425	3.334
Student Grade Level x Total PQA Score x SPP Status	–0.011	0.089
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.		

SOURCE: 21st CCLC Evaluation; PQA and APT-O observation scores from 49 activities; 401 youth engagement surveys; and TEA, PEIMS, 2013–14.

When considering results related to both observed engagement (derived from OCES scores) and youth-reported engagement (obtained from the youth surveys), there are a few findings that are consistent across both approaches to measuring engagement. In three of the four models exploring main effects, the SPP status of the center was related negatively to youth engagement in activities. Although it is not possible to ascertain definitively why this was case, it is hypothesized that the lower level of youth engagement may be related to the overt academic content covered in SPP activity sessions that largely targeted students who were especially academically at risk. Youth

already struggling academically (who may have a lower academic self-efficacy) may find activities with overt academic content less engaging than activities with less of a content focus.

The other consistent finding across two of three models employing this predictor was that provision of a supportive, interactive, and engaging learning environment as measured by the PQA was associated positively with higher levels of youth engagement. The only model in which this was not found to be the case was when engagement was examined specifically for students in Grades K–3. This K–3 finding also was not unexpected because many practices embedded in the PQA are meant to foster a sense of agency and belonging that resonate in a more significant fashion with older youth.

Finally, in order to explore the role individual practices specified in the PQA may have in supporting youth engagement, the evaluation team ran another six HLM analyses in which each of the six PQA subscales referenced previously were included alternatively as the sole activity-level predictor. The goal of these analyses was to explore which specific subscales were related especially to youth-reported levels of engagement.[26] Both student grade level and the SPP status of the center were included in these models as well. Only *Planning and Choice* was related significantly and positively to youth-reported engagement ($p < 0.10$). The fact that this was the only subscale related to youth-reported engagement when examined individually may suggest that the various practices represented in the PQA have a cumulative effect on supporting the cultivation of youth engagement. In this sense, it may be appropriate to focus efforts on providing learning environments that are characterized by the various practices and approaches detailed in the full tool.

Engagement Practices

The previous quantitative analysis is important for understanding how quality instructional practices were connected to student engagement in SPP and ACE-only programming. In order to understand further those practices that led to high levels of student engagement, the evaluation team also conducted qualitative analysis of data from the site visits that took place during both years of the evaluation. The site visits described elsewhere in this report helped the evaluation team to identify several key practices that appeared to be associated with high quality. During these site visits, the evaluation team observed programming and interviewed activity leaders and site coordinators. These data then were analyzed to highlight the staffing, instructional practices, and key components of each center. These key findings from 2012–13 and 2013–14 outlined here provide additional context for the quantitative information described earlier in the chapter and present important information for potential program improvement.

[26] These subscales are made up of the individual items or practices on the PQA that were observed during site visits in spring 2014 (the process for developing these subscales is described in detail in Appendix D).

- SPP and ACE-only activities used a variety of strategies to engage youth in programming. During activity leader and site coordinator interviews, the most commonly cited strategies were as follows:
  - Use of games and other "fun" and hands-on activities
  - Use of grouping strategies, including small groups, one-on-one time between activity leaders and students, and group work during which students can be with their friends
  - Offering choice in terms of types of enrichment activities or activities within academic enrichment
  - Use of differentiation—that is, personalizing learning to focus on the specific skills that students need help improving
  - Use of technology such as iPads, Kindles, and interactive whiteboards
- Although the analysis described earlier showed that ACE-only programs had high levels of engagement compared with that in SPP programs, interviews and site observations uncovered several ways in which SPP programs may have been able to engage students if practices were enacted in a manner consistent with the other findings in this chapter, including the following:
  - *More use of technology to support learning.* As noted previously, when combined with face-to-face instruction, computer-based learning can be an asset to student engagement. SPP programs tended to use more computer instruction than did ACE-only programs.
  - *Smaller group sizes.* SPP programs tended to have a smaller adult-to-youth ratio, which allowed activity leaders to differentiate instruction more and personalize learning so that students could work on improving specific skills. This finding is notable because observed youth engagement in the previously described analyses was associated with lower staff-to-student ratios.
  - *Less use of homework help and more individualized instruction incorporating choice and reflection.* ACE-only students tended to spend more time in homework help than did SPP students, who tended to receive more focused instruction designed to improve mathematics and reading skills. Although the previous analysis showed that, under some conditions, high mathematics and reading content were associated with lower engagement, SPP activities were also more likely to incorporate the quality practices associated with engagement, such as offering youth opportunities for choice and reflection, than were homework help and tutoring sessions.

These findings suggest that both SPP and ACE-only programs may be able to support greater levels of engagement through employing some of the key strategies described previously. Several of the engagement strategies that came out of the interviews are also practices that would lead to high scores on the Youth PQA and are those identified as being associated with high levels of youth engagement in the previously described analyses.

Conclusions and Key Findings

As the theory of change indicates, student engagement is connected directly to the experiences students have in quality OST activities. This chapter attempts to understand that connection better and answer, at least in part, three RQs:

- **RQ3:** How does the quality of delivery differ between SPP and ACE-only programming?
- **RQ4:** What instructional and administrative practices lead to high student engagement?
- **RQ5:** How engaged are young people in SPP and ACE-only programming? What is the relationship between quality of delivery and student engagement?

Analyses of data from site visits and program observations conducted during 2012–13 and 2013–14 revealed several key findings as follows:

- **A combined (face-to-face and computer-based) mode of delivery is associated with higher quality and youth engagement practices.** PQA scores tended to be highest for activities delivered in a combined format, particularly on the engagement portion of the PQA observation tool. This finding may indicate that a combined format is more conducive to youth engagement.
- **The *Learning Strategies* approach is associated with higher quality and youth engagement practices.** Again, PQA scores, particularly scores in the engagement portion of the tool, tended to be higher for activities delivered with a *Learning Strategies* approach, when they had a defined curriculum, than with a *Skills-Building* approach.
- **Low staff-to-youth ratio is associated with higher levels of observed youth engagement.** Activities that have lower staff-to-youth ratios had higher levels of observed youth engagement.
- **Practices outlined in the PQA are associated with higher levels of youth engagement.** In two of the three models exploring the relationship between PQA scores and youth engagement (both observed and self-reported), greater adoption of practices represented in the PQA was associated with higher levels of youth engagement in programming. This result was found when the total PQA score was used in analyses. Efforts to isolate which particular subscales of the tool may be related especially to engagement were less revealing. It seems that the individual practices outlined in the PQA may have an additive effect when present to support youth engagement in programming.
- **Youth engagement was lower in SPP programs.** In three of the four models exploring main effects, the SPP status of the center was related negatively to youth engagement in activities. Although it is not possible to ascertain definitively why this was the case, it is hypothesized that the lower level of youth engagement may be related to the overt academic content covered in SPP activity sessions that largely targeted students who were especially academically

at risk. Negative relationships between participation in academic enrichment activities and activities with high reading content further support this conclusion.

Generally, these findings support the theory of change indicating that higher quality activities are more likely to engage youth, and lower quality programs are less likely to engage youth. However, the fact that the relationship between program quality as measured by the PQA and youth engagement was weaker in SPP programs may indicate that the manner in which academic content is being delivered in SPP settings depressed youth engagement, even when positive youth development approaches described in the PQA were employed. Although it is unreasonable to expect that youth may be less engaged in programming with explicit academic content, it is still important to note this for program-planning purposes.

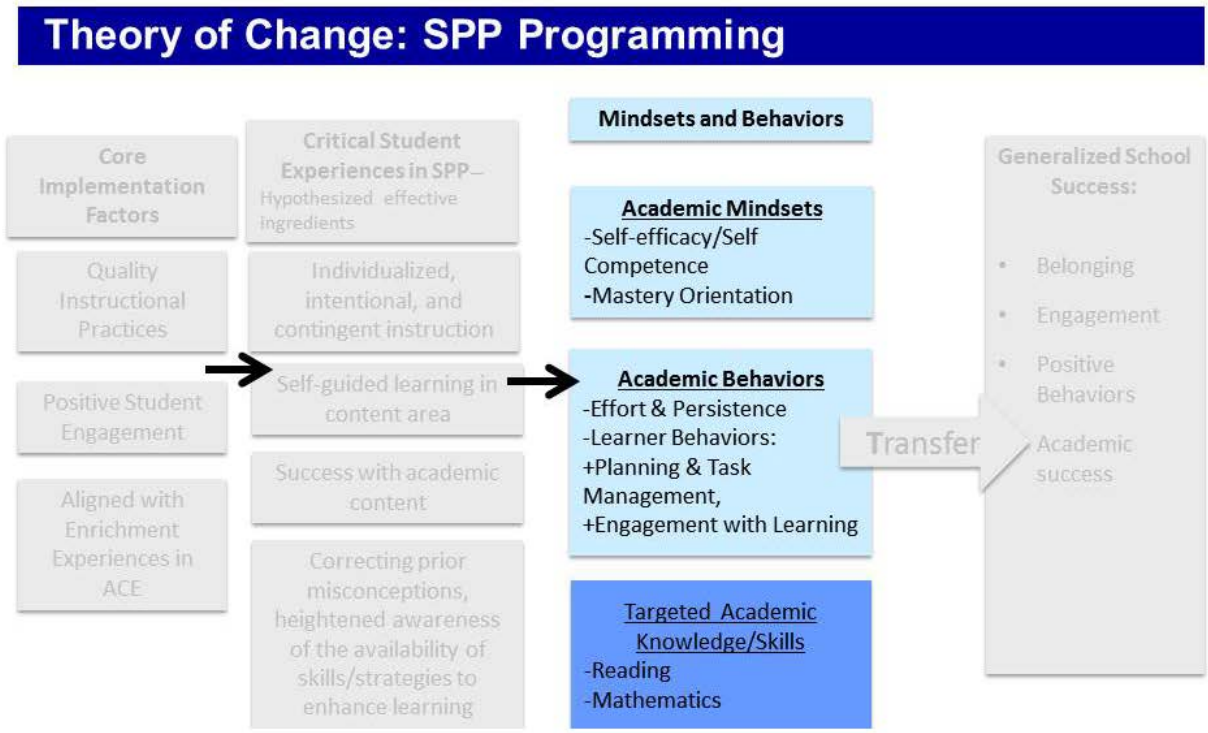
As noted previously, the theory of change posits and research supports that high levels of engagement are associated with improvement on a variety of student outcomes. Hence, it is important for both SPP and ACE-only activities to engage youth at high levels. According to this study, programs are able to engage youth better when the ratio between staff and youth is low, when academic content is less explicit (but not necessarily absent), when program quality is high, and when certain youth development practices are present. Therefore, professional development and training may be well advised to focus on these practices. It would be beneficial for TEA to consider planning professional development and training for both ACE-only and SPP activity leaders and site coordinators that builds on the findings from this analysis and focuses on quality practices outlined in the PQA, such as providing opportunities for youth choice and reflection, as well as ways to lower the ratio, embed academic content in less explicit ways, and build in opportunities for youth choice and reflection.

The next chapter examines student outcomes and, in doing so, considers the type of activities in which students participated. According to the theory of change, outcomes follow engagement in quality experiences, and, among the activity types in which students were most engaged, there may be positive change on certain measures.

Chapter 5: Initial Exploration of SPP Impact

The SPP theory of change illustrates how students might benefit from their participation in the SPP programs. In particular, the third column shows expected changes in student academic mindsets and behaviors that might result from high-quality instructional and organizational practices. As noted in the introductory chapter, academic mindsets and behaviors, such as persistence; a sense of self–competence; an ability to plan, manage, and engage with learning; and an orientation toward mastery, are associated with more generalized school success, such as improved grades and attendance and more engagement in learning (Farrington et al., 2012).

Figure 6. SPP Theory of Change, Chapter 5



Chapter 5 begins to explore this connection by examining the following RQ:

- **RQ6:** What is the impact of SPP programming on students' academic mindsets and behaviors? How does this compare with ACE-only students?

To answer this question, both SPP and ACE-only students were given a Mindsets and Behaviors Survey in 2012–13 and 2013–14. Students in Grades 4–12 received their own survey to complete (see Appendix E, Tables E7–E12 for specific survey items). Students in Grades K–3, who were considered too young to provide reliable responses on the Mindsets and Behaviors Survey, had an activity leader complete the survey on

their behalf.[27] Four types of analyses were conducted on the data. In both years, the evaluation team first conducted a psychometric analysis of the Mindsets and Behaviors Survey and Activity Leader Reports, which showed how well the surveys functioned—that is, how well the surveys measured what they intended to measure. Next, the team conducted an analysis of how SPP students' scores compared with ACE-only students' scores at the outset of the program, or at baseline, in each year. This method allowed the team to identify any initial differences to explore with the next set of analyses, a comparison of pre-post scores on surveys for both types of students to assess the impact of SPP participation on student outcomes. Finally, in 2012–13 only, the team compared SPP participants engaged in two different types of programming (those with a *Learning Strategies* approach and those with a *Skills-Building* approach) to learn which type of programming may be more effective in producing the desired changes in academic mindsets and behaviors.[28]

Analysis of the Functioning of Student Surveys

In 2012–13, the evaluation team conducted psychometric analyses of the two surveys (the Grade 4–12 student survey and the Grade K–3 activity leader report) to determine whether they were dependable measures of academic mindsets and behaviors. For a detailed description of these analyses, see Appendix H. Results showed that changes were warranted on both versions of the survey to make them more effective before being used in the 2013–14 evaluation. The evaluation team made two key changes to the surveys on the basis of these findings. The first major change was the inclusion of a new scale on the Grade K–3 survey focused on academic performance to be completed by teachers. This new section was designed to capture more detailed information about how students were performing in mathematics and reading to get a better understanding of the impact SPP programs might have had on the development of key academic skills. In some cases, the teachers completing this section of the survey were school-day teachers who did not teach in the OST program, and in other cases, the teacher was someone working in the OST program but familiar with the student's academic abilities. A second key change was to remove two scales completely from the Grade 4–12 survey—*Self-Efficacy* and *Mastery Orientation*. These were replaced by new scales focused on students' *Sense of Competence as a Learner*, *Sense of Competence in Reading*, and *Sense of Competence in Mathematics*. The team performed a second set of analyses in 2013–14 to test the functioning of the new versions of the surveys. In both cases, the new versions of the survey were improvements over the previous versions. Because the survey was revised between

[27] These surveys were developed specifically to meet the needs of this evaluation but are based on reliable and valid existing surveys including the Survey of Academic and Youth Outcomes. For more information on the Survey of Academic and Youth Outcomes, see http://www.niost.org/Training-Descriptions/survey-of-afterschool-youth-outcomes-youth-survey-sayo-y.

[28] The findings in this chapter are considered exploratory only. The 2012–13 data reflect pre-post data that were only a few months apart and were collected after the SPP programming had been implemented for only one semester. Data from 2013–14 are more robust, as they were collected after a full year of implementation and represent a greater spread of time between pretests and posttests, but the analysis presented from 2013–14 is only preliminary. A more statistically rigorous impact analysis will take place in the 2014–15 evaluation. The results presented here are used to inform areas for additional analysis.

2012–13 and 2013–14 and is markedly different, and because the analysis performed in the two years of the evaluation are different, results from each year are presented separately and chronologically.

Academic Mindsets and Behaviors—Comparing SPP and ACE-Only Students at Baseline

To begin to understand the impact of SPP participation on academic mindsets and behaviors, a first important step was to see how different SPP and ACE-only students were at the outset of programming, or baseline. Looking at differences at this stage both in 2012–13 and 2013–14 helped the evaluation team to understand the starting mindsets and behaviors of the two different types of students before looking at how much they changed on their responses on the survey by the end of the school year.[29,30] Because the SPP students were recruited based on their academic needs, one hypothesis was that their mindsets and behaviors related to academic performance would be different from those of students who participated exclusively in ACE-only programming.

After conducting a thorough set of analyses on the baseline data for both types of students across both years of the evaluation, we found it evident that some differences existed between the ACE-only and SPP students in their academic mindsets and behaviors at the time programming began, but the differences were not large. The full set of analyses and results can be found in Appendix I. These results may indicate that SPP students have certain academic-related mindsets and behaviors to a lesser extent than do ACE-only students. As noted earlier, this may be because SPP programs target students at risk for academic failure (and who are, therefore, potentially less academically capable or confident than are ACE-only students). On the other hand, limitations to the survey and survey administration also may explain the differences between the two groups of students. As noted earlier, many ACE-only students in the 2012–13 evaluation had participated in OST activities longer than the SPP students had, which means they may have had more exposure to tutoring, homework help, and

[29] Year 2012–13 was somewhat different from 2013–14 in that, during 2012–13, the ACE-only students had been involved in the ACE program since the fall 2012 semester, whereas many SPP students had begun receiving SPP programming much later in the school year during the spring 2013 semester. Although the survey was administered to SPP students (or the activity leaders who answered on behalf of the students) at the beginning of SPP programming, this was not the case for ACE-only students in 2012–13; ACE-only students were administered the survey in the middle of the programming year. SPP students at some centers participated in ACE activities in fall 2012, but many participated in ACE activities only after the SPP programs already had been started. All of this makes comparisons between the years difficult at best.

[30] In 2012–13, only SPP and ACE-only participants who attended at least *10 hours* of SPP or ACE-only program activities were selected for inclusion in the analyses for which results are described in this section of the report. In the second year of this evaluation, the analyses included students who had participated in ACE-only programs for a year, with a minimum of 30 days of program participation. However, because the SPP program had been in operation for less than a semester in 2012–13, a 10-hour minimum threshold was determined to be a more viable cut point for identifying a regular SPP student participant. In 2013–14, because there was a full year's worth of participation between preanalysis and postanalysis, the evaluation team reverted to including in the analyses students with 30 days or more of program participation.

other enrichment programming that supports the development of academic mindsets and behaviors. This finding does not explain the 2013–14 results, however, because in that year students took the survey at the same time, indicating that survey administration cannot account fully for observed differences.

Although the exact reason for differences between SPP and ACE-only at baseline is unknown, the value of this analysis was to uncover what differences may exist and highlight areas for future study in the 2014–15 analysis. In particular, the future study will focus on the pre-post differences in SPP students with ACE-only students and whether the intervention appears to minimize some of the differences found between the two types of students at the program's outset.

Academic Mindsets and Behaviors—Examining the Impact of SPP Programming Across Time

2012–13 Analysis

To examine the impact of SPP participation, the evaluation team assessed differences between the pre-post administrations. Specifically, the analysis compared the performance of students who participated in SPP with that of similar students who participated in ACE-only program activities. In order to create a comparison group of ACE-only students who were well matched to the SPP student sample, the evaluation team employed a statistical technique called propensity score matching (PSM) that allows researchers to match a sample of participants with nonparticipants who are similar on key characteristics related to the program outcomes (e.g., STAAR scores, special education status, English proficiency). The goal of this technique is to create two groups of students that are similar in every way except for their participation in SPP or ACE-only programming to allow for the strongest comparison (for more details on the PSM method, see Appendix D).

Once the sample was identified, the team conducted a comparative analysis (paired *t*-test) between the mean difference in scores from pre- and posttesting on the Mindsets and Behaviors Survey for SPP students and the ACE-only comparison group. Table 16 shows the effect of SPP programming on the key areas or constructs measured in the Grades K–3 and 4–12 surveys. Findings suggest that SPP programming had a positive impact on items related to *Effort and Persistence* and *Learner Behaviors* for students in both grade ranges. The Grade K–3 youth enrolled in the SPP program received an average of 3.2 points higher on the *Effort and Persistence* scale on the posttest and 3.8 points higher on the *Learner Behaviors* scale than did the comparison group, and these differences were statistically significant. Although the effect sizes for both constructs were small, they still suggest that SPP students had slightly more growth or positive change in their mindsets and behaviors during the course of the program than did their ACE-only peers.[31] However, these findings should be viewed with some degree

[31] Effect sizes can be difficult to interpret and depend on the field of study. One common interpretation is that 0.2 is considered a small effect; 0.5, a medium effect; and 0.8, a large effect (Cohen, 1988). However, effect sizes can be much smaller and still considered substantial. For example, an oft-cited

of caution. Because ratings were provided directly by the OST staff working with the students in the SPP programs, they are, in some ways, rating their own success in helping students to improve.

For students in Grades 4–12, findings were similar. It appears that students participating in SPP programming had slightly more growth in their mindsets and behaviors during the course of the program than did ACE-only participants for items related to *Effort and Persistence* and *Learner Behaviors and Engagement*. In both cases, the impact was statistically significant. Again, the effect sizes were small but suggest that SPP programming appeared to be having a slightly greater impact on academic mindsets and behaviors than did ACE-only programming.

Table 16. Impact of SPP Programming on Mindsets and Behaviors Constructs (2012–13)

Grade Level	Constructs	Effect Coefficient	SE[1] of Effect Coefficient	Effect Size	*p* Value
K–3	*Effort and Persistence****	3.227	0.713	0.106	< 0.0001
	*Learner Behaviors****	3.767	0.761	0.123	< 0.0001
4–12	*Self-Efficacy/Self-Competence*	–0.424	0.688	–0.041	0.538
	Mastery Orientation	1.021	0.662	0.101	0.124
	*Effort and Persistence****	1.713	0.663	0.167	0.0099
	*Learner Behaviors and Engagement**	1.102	0.664	0.108	0.097

[1] Standard error.
**p* < 0.10, ***p* < 0.05, ****p* < 0.01.
N = 1,794.

SOURCE: 21st CCLC Evaluation Activity Leader and Teacher Reports (Grades K–3) and Youth Mindsets and Behaviors Survey (Grades 4–12).

2013–14 Analysis

In 2013–14, the evaluation team used simpler analytic techniques to review the pre-post scores on the Mindsets and Behaviors Survey. Knowing that in 2014–15 the team will use rigorous techniques to examine the data closely, we designed this preliminary analysis in 2013–14 simply to identify descriptive pre-post trends and notable areas for future study.

In order to determine whether there were differences between SPP and ACE-only students in terms of how much they changed from fall 2013 to spring 2014 on the

---

study of the impact of aspirin on heart disease that transformed prevention of the disease showed an effect size of only 0.03. The effect sizes found in this study are small but are still consistent and substantial enough to conclude that a relationship exists and further study is warranted.

Mindsets and Behaviors survey, the evaluation team conducted analyses of the change from pre- to posttest on the survey for all SPP and ACE-only students who had taken a survey at both time points. Table 17 shows the results of this analysis.

For Grades K–3 (for which ratings were provided by OST staff on behalf of students), both SPP and ACE-only students witnessed significant improvement (analysis was done using a paired *t*-test) from the fall to the spring on all three of the constructs that were measured by the survey—*Effort and Persistence*, *Learner Behaviors*, and *Academic Performance*. This follows what is posited in the SPP theory of change—that students will improve their key mindsets and behaviors on participation in high-quality programming. The analysis in Chapter 4 shows that SPP and ACE-only programs were demonstrating average or higher levels of quality, so according to the theory we would expect moderate growth in students' mindsets and behaviors. Although there was growth, in all three cases the ACE-only students saw the same or higher levels of growth as the SPP students. Given the preliminary nature of the analysis, there is no way to know why ACE-only students may have demonstrated more growth or whether this growth will hold across time. The improvement in both groups is a positive sign and worth exploring further in 2014–15 to understand whether the differences between the two groups persist and if growth continues.

For students in Grades 4–12, the results were quite different. For SPP students, the only area that showed significant difference from fall to spring was the *Sense of Competence as a Learner* scale. However, the difference was a *decline* from fall to spring. Likewise, ACE-only students saw significant differences from fall to spring on all five constructs, but in all cases, the difference represented a decline from fall to spring. There is not enough information at this time to know what is causing those declines, and there could be inherent differences between the two types of students that might explain the differences. There are also challenges inherent with any pre-post survey in which students rate themselves. Called *response shift bias*, the phenomenon happens when students have one frame of reference at the start of a program when filling out a survey and a different frame of reference when completing a survey at the end of the program. They are rating themselves on the basis of different knowledge or experiences and therefore are not able to represent change across time accurately (Howard, 1980).

Table 17. Pre- and Posttest Mean Differences on Mindsets and Behaviors Constructs (2013–14)

Grade Level	Constructs	SPP Students				ACE-Only Students			
		Fall 2013 Mean	Spring 2014 Mean	Mean Difference	N[1]	Fall 2013 Mean	Spring 2014 Mean	Mean Difference	N[1]
K–3	*Effort and Persistence*	49.1	50.4	1.3**	373	47.9	51.2	3.3***	252
	Learner Behaviors	48.8	50.7	1.9***	373	47.8	50.9	3.1***	251
	Academic Performance	45.7	49.0	3.2***	334	52.0	55.3	3.2***	276
4–12	*Sense of Competence as a Learner*	50.9	49.7	–1.3**	505	51.2	49.1	–2.0***	389
	Sense of Competence as a Reader	50.2	49.6	–0.6	505	51.1	50.3	–0.8	389
	Sense of Competence in Mathematics	51.0	50.4	–0.7	507	51.2	49.5	–1.7***	390
	Effort and Persistence	50.9	50.2	–0.7	507	52.0	49.4	–2.6***	390
	Learner Behaviors and Engagement	50.6	50.0	–0.6	507	52.1	49.1	–3.0***	390

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$ (within-group comparison).

[1] These N counts represent students who had both a pretest and a posttest, so the numbers differ from the numbers at baseline noted elsewhere.

SOURCE: 21st CCLC Evaluation Activity Leader and Teacher Reports (Grades K–3) and Youth Mindsets and Behaviors Survey (Grades 4–12) (2013–14).

When reviewing Table 17, it is important to keep in mind that these differences cannot be attributed definitively to any particular cause; by the nature of the statistical test used for the analysis, all that can be concluded here is that there are statistically significant differences between scores (fall to spring) within the two groups. The impact analysis in 2014–15 using stronger statistical controls and a matched sample should help to uncover whether any of these effects are attributable to the program or whether other factors (unknown or merely theorized) are the cause of these differences.

Academic Mindsets and Behaviors—Examining the Relationship Between Program Types and Youth Outcomes

One purpose of the SPP evaluation is to learn whether certain types of programs and practices more than others generate positive outcomes for student participants. As described earlier in Chapter 3, in 2012–13, the evaluation team created program typologies based on the general learning approach applied in the program. Two basic program types were identified: those that used a *Learning Strategies* approach, which helps students develop skills that can be transferred to different subject areas, and those that primarily used a *Skills-Building* approach, which focuses more on specific content-area skills.

The hypothesis was that the students who participated in programs using primarily a *Learning Strategies* approach would have better outcomes (e.g., more improvement on the Mindsets and Behaviors Survey or on their STAAR scores) than would students in programs that primarily used a *Skills-Building* approach. One reason for this was that the observed *Learning Strategies* activities emphasized application of knowledge and skills, self-assessment, metacognition, and so forth—skills that might transfer to other subject areas. In comparison, the *Skills-Building* activities were more focused on specific, discrete skills related to the content area. Another reason was that, for the most part, the observed activities using a *Learning Strategies* approach generated higher levels of student engagement than did other activities.

In order to examine the question of which type of program produced better outcomes, specifically for students in Grades K–3,[32] the evaluation team used an HLM approach (see Appendix D for details on the method). Two constructs, *Effort and Persistence* and *Learner Behaviors*, were examined. The analysis was aimed simply at determining whether there was a relationship between participation in a certain type of program and better outcomes but was not attempting to prove that the type of program was the cause for the change. In other words, the analysis was correlational, not causal. The analysis controlled for students' prior performance (on the pretest) on the two constructs and other covariates including gender, ethnicity, economically disadvantaged status, limited English proficiency status, number of years enrolled in the program, and total days of SPP program participation.

As Table 18 shows, students participating in SPP programs with a *Learning Strategies* approach had significantly more growth on the items in the *Effort and Persistence* construct than did students participating in programs using the *Skills-Building* approach. Specifically, students attending programs using a *Learning Strategies* approach scored almost 6 points (on a 0–100 scale) higher on the posttest than did those enrolled in SPP programs using a *Skills-Building* approach. A moderately significant impact also was found for the construct *Learner Behaviors.* Students attending programs associated with the *Learning Strategies* type of SPP programs scored a little more than 5 points higher on those items than did students enrolled in SPP programs that used a *Skills-Building* approach.

[32] The sample size was too small for students in Grades 4–12 to include them in the analysis.

Table 18. Comparative Analysis of Impact of SPP Programming Using a *Learning Strategies* Approach Versus a *Skills-Building* Approach on Mindsets and Behaviors (2012–13)

Grade Level	Constructs	Effect Coefficient	SE[1] of Effect Coefficient	*p* Value
K–3	*Effort and Persistence***	5.951	2.816	0.035
	*Learner Behaviors**	5.268	3.008	0.081
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. [1] Standard error. *N* = 405				

SOURCE: 21st CCLC Evaluation Activity Leader and Teacher Reports (Grades K–3).

The findings from this analysis indicate that the hypothesis—that the *Learning Strategies* approach may be more effective in supporting the growth of certain mindsets and behaviors than is the *Skills-Building* approach—is valid. Because programs using a *Learning Strategies* approach focus on broad skills and how to make learning applicable across content areas, they may be able better to foster some of the skills measured by the survey than are other types of programs.

Conclusions and Key Findings

This chapter attempts to answer the following RQ:

- **RQ6:** What is the impact of SPP programming on students' academic mindsets and behaviors? How does this compare with ACE-only students?

The findings from the analysis of the Academic Mindsets and Behaviors Survey in 2012–13 and 2013–14 are mixed, but overall they provide some evidence to support the SPP theory of change as indicated by the following key findings:

- **SPP programs appear to be recruiting their intended student population.** The analysis of student responses on the Mindsets and Behaviors Survey at the program outset for both 2012–13 and 2013–14 indicates that SPP students are somewhat less skilled and confident in their mindsets and behaviors than are their ACE-only peers. These differences were particularly strong in the areas of *Learner Behaviors* (Grades K–3 and 4–12)*, Effort and Persistence* (Grades K–3)*,* and *Academic Performance* (Grades K–3).
- **There appears to be a small but positive impact of SPP programming on mindsets and behaviors.** The 2012–13 results (impact analysis) showed a small, positive impact of SPP programming on *Effort and Persistence* and *Learner Behaviors* for Grades K–3 and a positive, small impact on *Effort and Persistence* for Grades 4–12. These findings used statistical techniques to ensure the comparison between groups was valid and to eliminate as much bias as possible that could affect the findings. This positive finding in the first year of the program highlighted a positive trend worth exploring further in data from the second year of the program. Because the program was so new and there was a

short window of time between pretest and posttest, the evaluation team wanted to review results from the second, full year of programming to see whether the positive effects from 2012–13 held.

In 2013–14, the findings were not as conclusive. The analyses during this year were preliminary and descriptive in nature to help highlight trends in preparation for a more rigorous impact analyses in 2014–15. The analyses considered pre-post change in scores on the survey for both types of students and found that there was a statistically significant, positive change between fall and spring for the SPP and ACE-only Grade K–3 group for all three constructs tested, with the changes being larger for the ACE-only group. For Grades 4–12, there were no statistically significant differences in mean scale score for any construct in the SPP group, whereas the ACE-only group displayed statistically significant mean scale score differences for which the spring score was actually lower than the fall score. At this juncture, extreme caution should be exercised in drawing any substantive conclusions from these findings. Comparisons between SPP and ACE-only students on these constructs are premature at this point because steps have not been taken to create a matched sample of ACE-only students who are similar to SPP students. Furthermore, efforts to control for the participation profile of youth and organizational and POS quality have not been introduced into the models yet to explore how programming may impact youth on these outcomes when both quality and engagement are high. Each of these questions will be explored more thoroughly by using impact analysis in 2014–15.

- **A *Learning Strategies* approach appears best suited to foster improvement in mindsets and behaviors.** The 2012–13 analyses that assessed differences in program impact considered whether an SPP program primarily used a *Learning Strategies* or a *Skills-Building* approach. Results suggest that Grades K–3 youth participating in SPP programming using a *Learning Strategies* approach demonstrated a higher level of improvement on both the *Effort and Persistence* and *Learning Behaviors* scales (Table 18). There are a few important caveats to these findings—namely, that the findings are still preliminary and are primarily correlational and descriptive in nature. The analyses cannot answer the question of whether enrollment in programs adopting a *Learning Strategies* approach caused youth to improve to a greater degree on the *Effort and Persistence* and *Learner Behavior* scales. It is possible that this correlational finding can be explained by other unobservable characteristics of participating students or that the level of quality, which was observed to be higher among programs in the *Learning Strategies* group, drove the positive findings. The relationship between program typologies and outcomes will warrant additional study in 2014–15, Year 5, of the evaluation.

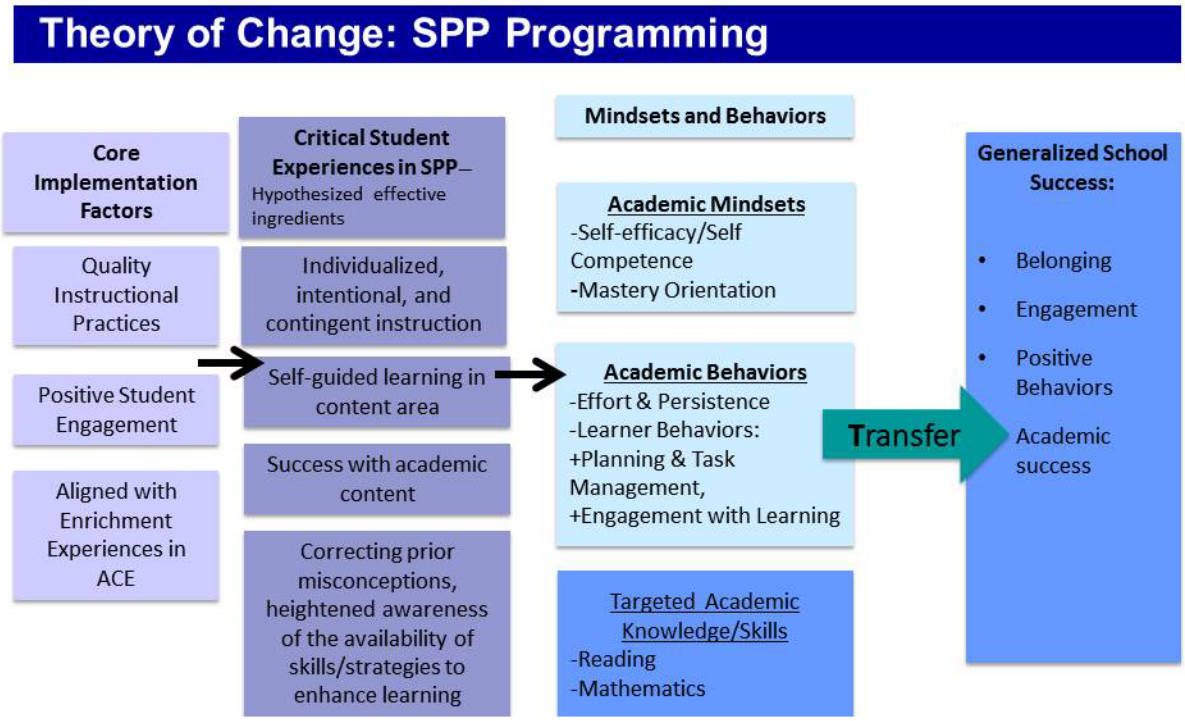
Chapter 6: Conclusions and Next Steps

This report combines findings from 2012–13 and 2013–14 of the SPP program. As noted elsewhere in the report, the two years' worth of data collection involved different students and staff, so it is not longitudinal in nature. However, taken together, the data from both years tell an important story about the initial implementation of the SPP program and the initial impact it may have had on students who participated. In this chapter, we summarize the preliminary findings from the two years of evaluation and make some preliminary conclusions about the SPP program. We also make recommendations for how these findings might impact the 2014–15 evaluation activities as well as program implementation and planning activities.

A Summary of Preliminary Findings

As noted throughout this report, the SPP program was based on a theory of change positing that participation in high-quality experiences can lead to improved mindsets and behaviors in youth participants, which in turn can transfer into generalized school success. The evaluation explores this theory of change by examining program quality, staff experiences in the program, youth mindsets and behaviors, and school-related outcomes.

Figure 7. SPP Theory of Change, Conclusion



The evaluation was organized around six RQs associated with implementation and impact. The preliminary findings across RQs are presented in this section. Some RQs were addressed only partially at this time because of the intention to conduct more rigorous impact analyses in 2014–15. The 2012–13 and 2013–14 evaluations were intended to be more descriptive and exploratory in nature in order to inform additional analyses in 2014–15.

RQ 1. How does SPP programming compare with ACE-only programming in centers administered by the same grantee?

Overall, findings from 2012–13 and 2013–14 indicate that **SPP programs are more academic in nature than are ACE-only programs.** This finding was expected, given the nature of the funding for SPP programming and the stated and explicit intention that SPP programs would help students at risk for academic failure improve their skills. Specific findings that contribute to this overall conclusion include the following:

- *SPP programs hired more certified teachers.* In both years of study, a far higher percentage of SPP activity leaders (85 percent in 2012–13 and 81 percent in 2013–14) than ACE-only activity leaders (51 percent in 2012–13 and 42 percent in 2013–14) were credentialed teachers.
- *SPP students spent more time in academic activities. S*tudents participating in SPP activities spent much more of their time in academic activities—particularly in academic enrichment activities and tutoring (82 percent in 2012–13 and 66 percent in 2013–14 for SPP students compared with 45 percent in 2012–13 and 33 percent in 2013–14 for ACE-only students). Compared with ACE-only students, SPP students spent very little time in recreational activities (8 percent versus 33 percent in 2012–13 and 25 percent versus 33 percent in 2013–14). Students participating in SPP activities also spent much less of their time in homework help sessions than did students participating in ACE-only activities. In 2013–14, SPP students began participating in recreation programs and other nonacademic enrichment to a larger degree, but there was still a substantial difference between the participation of the two types of students.
- *SPP activities used smaller groups and longer activities to support academic learning.* SPP activities typically met in smaller groups than did ACE-only activities, and SPP activity leaders served fewer students (on average six fewer) per week than did ACE-only activity leaders. In addition, on average, SPP academic activities were 90 minutes long, compared with 60 minutes for ACE-only academic activities.
- *SPP and ACE-only activity leaders used different instructional strategies.* SPP activity leaders were more likely than ACE-only activity leaders to plan activities that explicitly promoted skill building related to state standards. They were also more likely than ACE activity leaders to use instructional strategies that explicitly addressed content knowledge—the use of computer-based learning programs, direct instruction, and the review and practice of concepts learned during the school day.

RQ 2. How do students participating in SPP programming differ from students who participate in ACE-only programming?

Analysis of student assessment data and youth survey data revealed that the students participating in SPP programming tended to be more academically at risk and less proficient in key academic mindsets and behaviors[33] than were their ACE-only peers at program onset. This finding is important because it indicates that the SPP program was successful in recruiting the types of students it intended to serve—that is, those at risk for academic failure. Key differences between the two types of students were as follows:

- *SPP students have higher academic needs than do ACE-only students.* Students who attended SPP programs in 2012–13 appeared to have higher academic needs than did ACE-only students, as shown by their performance levels on the 2012 STAAR assessment. Specifically, between 19 percent and 36 percent fewer SPP students than ACE-only students met satisfactory proficiency levels on the various STAAR exams in 2012.
- *SPP students scored lower on the Mindsets and Behaviors Survey than did ACE-only students.* SPP students at all grade levels and across almost all constructs in both years scored lower than ACE-only students did on the Mindsets and Behaviors Survey at the program outset. In several cases, those differences were statistically significant. ACE-only students appeared, in particular, to be at a significantly higher level at the start of programming on constructs measured on the survey, including *Learner Behaviors* (at the K–3 and 4–12 grade levels) and *Academic Performance* (at the K–3 level).

RQ 3. How does the quality of delivery differ between SPP and ACE-only programming?

The evaluation team conducted analyses of the overall quality of both program types and found that SPP programs and ACE-only programs had roughly the same level of overall quality and that the level was average on the basis of observation ratings. The two types of programs were not the same, however, and the evaluation team found that activity leaders differed in their administrative practices in several key ways. Those differences, outlined next, may contribute to other findings in the report related to levels of youth engagement in the activities.

- *SPP activity leaders had more* Internal Communication*.* Activity leaders for SPP activities reported higher levels of *Internal Communication* than did those who led ACE-only program activities.
- *SPP activity leaders were more likely than ACE-only activity leaders to report* Program Challenges*.* The challenges they cited most frequently were inconsistent program attendance, student behavioral issues, and low levels of student interest in the activities. Challenges with technology were also present.

[33] Students were measured on the academic mindsets and behaviors through a survey given at program onset and again the spring after program completion. The survey measured competencies such as *Effort and Persistence*, *Learner Behaviors*, *Engagement*, and *Sense of Competence as a Learner*.

- *SPP activity leaders focused on academic skills.* SPP activity leaders were more likely than ACE-only activity leaders to report implementing programs and activities focused on specific academic skills, particularly those connected to the state standards, and to direct students in *Academic Learning Activities* more frequently than were ACE-only activity leaders.

These differences, although not directly related to the POS, affect the quality of program delivery. The last finding, in particular, is important because it connects to other findings in the report related to levels of youth engagement in explicitly academic programming.

RQ 4. What instructional and administrative practices lead to high student engagement?

By reviewing scores from observations of youth engagement, as well as youth responses to an engagement survey, and connecting those with observed and reported activity leader practices, the evaluation team was able to conclude that **activities that used a *Learning Strategies* approach, a combination of computer-based and face-to-face delivery, and a low staff-to-youth ratio, were the most engaging to young people.** The specific findings that supported this overall conclusion include the following:

- *Generalized learning strategies were associated with engagement.* SPP programs had two different approaches associated with curricular content. One was a *Learning Strategies* approach, which emphasized general learning strategies that were applicable across different content areas. A second common approach was a *Skills-Building* approach, which emphasized specific skills associated with a subject area. The *Learning Strategies* approach was associated with higher levels of quality than any other approach. That, combined with the findings in this report connecting quality and engagement, suggests that students may be more engaged when participating in activities using a *Learning Strategies* approach.
- *A combination of face-to-face and computer-based instruction was most engaging to students.* SPP activities also used two modes of delivery: computer-based delivery, in which students typically worked individually on an academic skills computer program, and face-to-face delivery, in which students worked mainly in small-group sessions facilitated by an activity leader. The majority of SPP programs used a combination of computer-based and face-to-face delivery modes. The observations of the SPP program activities, as well as the student engagement survey, indicated that engagement levels were higher among students in activities that combined computer-based and face-to-face delivery modes than among students in activities using only a single mode of delivery. Reported engagement was lowest among the students in only computer-based activities.
- *Low staff-to-youth ratio was associated with youth engagement.* A low staff-to-youth ratio was associated with higher levels of youth engagement during observations of youth engagement.

RQ 5. How engaged are young people in SPP and ACE-only programming? What is the relationship between quality of delivery and student engagement?

An analysis of 2013–14 quality and engagement data showed that although overall there was not a significant difference between the quality of ACE-only and SPP programs, **youth-reported engagement was lower in SPP programs than in ACE-only programs.** The analysis of these two types of data also revealed that, **overall, higher quality programs engaged students at higher levels, but in SPP programs, the relationship between program quality as measured by the PQA and youth-reported engagement was weaker, a finding that is hypothesized to be related to the provision of higher levels of academic content in SPP programs.** More specifically, the evaluation team found the following:

- *The* Learning Strategies *approach was associated with higher levels of quality than was any other approach.* The three SPP activities with the highest PQA scores used a *Learning Strategies* approach for which there was an established curriculum.
- *SPP programs were associated consistently with lower levels of youth-reported engagement.* It is hypothesized that this finding is related to the provision of higher levels of academic content in these programs.
- *Youth development principles were associated with youth engagement.* SPP and ACE-only program activities that employed youth development principles like those outlined in the PQA (e.g., engaging youth in activities that lead to a tangible product, offering youth choice, providing opportunities for reflection) had higher levels of youth self-reported engagement.

RQ 6. What is the impact of SPP programming on students' academic mindsets and behaviors? How does this compare with ACE-only students?

Findings related to this RQ were somewhat inconclusive, partly because the 2013–14 analysis was preliminary in nature at the time of this report. However, **there appears to be a small but positive impact of both SPP programming and potentially ACE-only programming on many of the measured academic mindsets and behaviors.** Again, this is important because improved mindsets and behaviors ultimately can lead to improved academic outcomes according to the SPP theory of change. Findings contributing to that conclusion include the following:

- *The 2012–13 results showed growth on most constructs of the Academic Mindsets and Behaviors Survey*. The 2012–13 impact analysis showed some significant growth for students participating in SPP programming on the academic mindsets and behaviors measured on the student survey, including *Effort and Persistence* and *Learner Behaviors* for Grades K–3 and on *Effort and Persistence* for Grades 4–12. However, the amount of time between pre- and postsurveys was very short because SPP programming ran for only one semester in 2012–13, so students took the pretest in January and the posttest in April.

- *A* Learning Strategies *approach may support improvement in academic mindsets and behaviors.* Results suggest that students in Grades K–3 who participated in SPP programming using a *Learning Strategies* approach demonstrated a significantly higher level of improvement on the Academic Mindsets and Behaviors Survey on *Effort and Persistence* and *Learning Behaviors* than did students in Grades K–3 who participated in SPP programming using a *Skills-Building* approach.
- *The 2013–14 results on the Academic Mindsets and Behaviors Survey were mixed.* In 2013–14, the analyses considered the pre-post change in scores on the Academic Mindsets and Behaviors Survey for both types of students. Changes were positive and significant from fall to spring for both SPP and ACE-only students in Grades K–3 on all of the areas measured by the survey. The changes from pre- to posttest were higher for ACE-only students, indicating that the program may have had a greater impact on ACE-only students' mindsets and behaviors than on SPP students'. For Grades 4–12, the analysis showed no improvement on the constructs for SPP students and a decrease in the mindsets and behaviors for ACE-only students—that is, their scores declined from fall to spring. These findings are inconclusive and should be viewed with extreme caution. The evaluation team has not yet employed control techniques to ensure a matched comparison group and has not performed the more rigorous analysis needed to answer some of the questions these preliminary findings raise. Currently, they simply point to an interesting finding worth exploring in more detail in the 2014–15 analysis.

Taken together, these findings begin to tell a story about the validity of the SPP theory of change, a story that will be completed in 2014–15 during the final year of this evaluation. Interestingly, the findings from the Years 3 and 4 analyses indicate that the theory of change may be applicable to both SPP and ACE-only programs—that is, overall, the 2012–13 and 2013–14 analyses confirm that SPP and ACE-only programs are delivering activities that collectively are of average quality but that range from average to high quality and are roughly the same across both program types. This is an important first step in the theory of change. Students also report engagement when youth development practices are present, when the delivery is both face-to-face and computer based, and when activity leaders used a generalized learning approach rather than specific skill building. This finding was true across both types of programming. The findings also suggest a connection between quality and engagement, the first relationship posited in the theory of change. This finding was true for programs overall but was stronger in ACE-only programs. Finally, the evaluation findings show some evidence that both SPP and ACE-only programming had an impact on the mindsets and behaviors of some participants. So far these effects are small (2012–13) and inconclusive (2013–14) but show enough positive trends to warrant further study, particularly to understand better the progression from quality programming to engaged youth to outcomes related to mindsets, behaviors, and school performance.

Next Steps

The findings described are important for two reasons: (1) the 2012–13 findings were used to improve on methodology and inform 2013–14 data collection, and (2) the 2013–14 findings will be used to inform planned analyses for 2014–15. Each of these data uses is described in greater detail in the following sections.

Methodological Changes to the Evaluation

The 2012–13 evaluation procedures provided valuable experience, on which the evaluation team drew to improve 2013–14 data collection, including the following:

- The evaluation was conducted during a very short time in spring 2013. During 2013–14, evaluation activities spanned the full school year to allow more time for change to occur among students and to ease the burden on sites by having collection periods spread out.
- Analysis of the validity and reliability of the student survey instruments revealed that two scales on the Grade 4–12 student survey were not functioning well. The evaluation team removed those scales from the survey. In particular, those scales were *Self-Efficacy/Self-Competence* and *Mastery Orientation*. Those scales were replaced with three new scales focused on *Sense of Competence as a Learner, Sense of Competence in Reading,* and *Sense of Competence in Mathematics.* In addition, the Grade K–3 survey had several items for which it was too easy for staff to rate the items highly. Those items were removed from the 2013–14 survey. Finally, a new scale was added to the Grade K–3 survey focused on academic performance and intended to be completed by school-day teachers familiar with the student's abilities in order to begin examining the perceived impact of the program on academic skills and to align with the addition of the *Sense of Competence* scales in the Grade 4–12 survey.

Likewise, findings from the 2013–14 evaluation will inform 2014–15 evaluation plans. For example, the evaluation team will use preliminary findings from 2013–14 about quality and engagement to inform planned analyses related to the SPP theory of change. In particular, using 2013–14 data, the team will continue to explore how quality influences student engagement and what the mediating effects of engagement and academic mindsets and behaviors are on key outcomes. Much of the data and information presented in this combined, two-year report are preliminary and offer a first look at the SPP program delivery, quality, students, and staff. In the coming year of the evaluation, the evaluation team will explore these findings in greater detail and conclude the exploration of the SPP theory of change begun with this report. Findings from that evaluation will be used to make recommendations about program implementation and professional development for staff.

Afterward: 2014–15 Evaluation Plan

This report has presented a wealth of information collected and analyzed in 2012–13 and 2013–14 related to program quality, youth engagement, and the exploration of the differences between SPP and ACE-only programs. As the evaluation moves into 2014–15, the evaluation team will use the data collected during 2013–14 to conduct additional analyses aimed at understanding the impact of SPP and ACE-only programs on youth participants and how that relates to the quality of programming and student engagement. These analyses, although focused explicitly on SPP programming, will contribute to the overall objectives of the five-year evaluation set out by TEA—to **(1) identify and describe innovative strategies and approaches** and **(2) conduct a statewide assessment of 21st CCLC programs.** The evaluation activities we propose for Year 5, broken out by objective, are as follows.

Identify and Describe Innovative Strategies and Approaches

Explore the relationship between SPP and ACE program quality and impact. Using the quality data obtained from observations and surveys collected in 2013–14, the evaluation team will conduct extensive analyses to test the SPP theory of change looking at how quality and engagement may lead to changes in skills and beliefs and, ultimately, improve school success by using more rigorous statistical techniques. Ultimately, the goal of this analysis is to explore whether there is evidence of a relationship between SPP and ACE program quality and outcomes, how this relationship varies between SPP and ACE-only sites, and the role of student engagement and changes in skills and beliefs mediating this process. This method will allow us to identify key practices, strategies, and approaches that are related especially to improving the quality of programming, leading to higher levels of youth engagement.

Explore changes in youth mindsets and behaviors. In 2014–15, the evaluation team will explore the findings from the Mindsets and Behaviors Survey, paying particular attention to the pre-post change from 2013–14 that showed some inconclusive results in the preliminary analysis. The goal of the 2014-15 analyses will be to control for differences between student populations and create a more accurate picture of what did and did not change for each type of student during the 2013–14 school year as a result of participating in SPP and ACE-only programs. This method also will explore the connections among quality, engagement, mindsets and behaviors, and other student outcomes as posited in the theory of change for SPP students in order to identify particular practices or approaches that might be leading to higher levels of change.

Conduct a Statewide Assessment of 21st CCLC Programs

Compare program impacts from SPP program participation relative to participation in ACE-only programs. Replicating impact analyses undertaken in 2011–12 accomplished through imposition of a quasi-experimental design using PSM, we will calculate estimates of program impact for the full domain of ACE-funded programs operating during the course of the 2013–14 school year. Such an effort will

involve comparing how students attending ACE programming at different levels (e.g., 30 and 60 days) performed on a series of academic outcomes (e.g., performance on STAAR exams, attendance in school) relative to a matched sample of students who did not participate in the program. Similar analyses will be run specifically for students served in the SPP program. This method will allow for a comparison of program effects between ACE-only and SPP programs operating in 2013–14. This latter analysis will include an impact analysis on change in youth mindsets and behaviors from the youth surveys collected in 2013–14 between SPP and ACE-only participants.

Explore significance and meaning of larger program effects in relation to high school students. One of the key findings from the 2011–12 evaluation was that program effects were greatest for high school students participating in ACE programs on several outcomes. What is not clear from these findings is whether some key, *nonobservable* attributes are associated with high school students opting to participate in 21st CCLC that makes them substantively different from students who choose not to participate in the program. In order to explore these issues further, in 2013–14 AIR developed a new survey instrument for use with a sample of ACE participants. The survey was pilot-tested with a small sample of ACE participants during spring 2014. A description of this initial effort is presented in the next section. In 2014–15, the evaluation team will administer youth surveys among the full population of ACE participants. When attendance data for the program become available in June 2015, the evaluation team will create three attendance groups—0–30 days, 30–60 days, and 60 or more days—and conduct analyses to see whether differences exist among the three groups on the survey. The goal of this data collection activity is to obtain data on a series of hypothesized factors that may differentiate regularly participating high school students from their peers who participate at lower levels in the program. These data then can be combined with the impact analyses described in evaluation activity two to explore the degree to which large program effects for high school remain after controlling for a broader array of student characteristics that might be related to participation in the program.

Assess SPP cost-effectiveness and sustainability. To determine the cost-effectiveness of the SPP programs, the evaluation team conducted a descriptive analysis of costs in 2013–14 of the evaluation. That preliminary study is presented in the afterward to the report. To conduct a full cost-effectiveness and sustainability study in 2014–15, it will be necessary first to identify the costs of the program on a per-student basis and then identify the impacts of the program (both positive and potentially negative) in terms of student outcomes (e.g., school-day attendance, performance on STAAR tests). Finally, the per-student costs relative to student outcomes will be determined to assess the cost-effectiveness of the SPP program. School-related outcome data will become available to the evaluation team in 2014–15, allowing the team to calculate program impacts and then tie per-pupil costs to those impacts. More specifically, the cost-effectiveness of the SPP program will be explored in several ways, as follows:

- To examine relationships between site-level per-pupil spending and student outcomes, site's cost-per-student ratios will be used in statistical models to predict site-level student outcomes, such as school-day attendance and STAAR

test scores.[34] These analyses will seek to identify whether cost-per-student ratios are related systematically to student outcomes, as well as the magnitude and direction of those relationships. For example, are schools with higher cost-per-student ratios significantly more associated with better student outcomes?

- Once the general relationship between site-level spending and outcomes is established, follow-up analyses will explore the extent to which spending on key program components and activities (e.g., payroll, face-to-face interventions) may provide further insight into the relationship between spending and outcomes. The rationale behind this strategy is that the manner in which centers allocate grant funds across program components may be related to how effective their program is for students participating in the SPP activities. The evaluation will assess the relationship between changes in the portion of funding spent on particular program activities and student outcomes associated with impact analyses carried out in 2014–15.
- A return-on-investment rating will be generated for each site. This rating will be an indicator of student outcomes achieved relative to per-pupil spending for each site compared with those of other SPP programs. This rating will be done by distributing sites into categories based on their relative ranking on outcomes (e.g., sites with best outcomes will be in the top category, sites with the lowest achievers in the bottom category, and so on) and on their relative ranking in terms of per-pupil spending, with the highest per-pupil spending sites in the top category, the lowest in the bottom category, and so on. Then sites' relative outcomes and spending will be placed into a matrix to display those sites' return on investment (see Table 19 for an example). The matrix will illustrate which sites achieve the best outcomes with the lowest per-pupil spending and vice versa. Per-pupil spending may be adjusted to ensure comparability across different geographic regions, site-level enrollment, and so on.

Table 19. Sample Spending Matrix

		Student Outcomes	
		High	**Low**
Per Pupil Spending	**High**	SPP site A—high spending, high outcomes	SPP site C—high spending, low outcomes
	Low	SPP site B—low spending, high outcomes	SPP site A—low spending, low outcomes

The analyses described previously will result in the 2014–15 final report that explores the impact and cost-effectiveness of SPP and ACE programming and will provide TEA with valuable conclusions to use in planning for the future of these important programs.

[34] The statistical models will include controls for factors that may impact student outcomes but are not related to SPP costs per pupil. These factors include, but are not limited to, site-level attendance and academic performance prior to the start of the SPP program.

Afterward: High School Study

As part of the 2013–14 evaluation of the 21st CCLC program, the evaluation team originally planned to develop a new high school survey for use in further exploring the large effect sizes found in the 2011–12 analyses. After discussion with TEA, it was determined that in order to ensure a better sample (i.e., capturing students' attitudes and perceptions at the beginning of the school year before they had participated in programming rather than at the end) and a more efficient data collection process, AIR would administer the high school survey in fall 2014 as part of the 2014–15 evaluation rather than 2013–14. However, the team developed the survey and piloted it during the spring of 2014 in order to test the items and determine the appropriateness and fit of the survey. The following sections outline the survey development process as well as the results of the small pilot conducted in May and June 2014 in three schools.

Survey Development

The survey was developed through a collaborative process between TEA and AIR. The survey draws on existing surveys developed by Policy Studies Associates, the Chicago Consortium on School Reform, and the National Institute on Out-of-School Time. Items are designed to understand students' academic motivations and beliefs, their future orientation, and their satisfaction with the ACE program. These survey results ultimately will be used to explore the previously unobservable characteristics of high school students who participate in ACE programs. The characteristics may help to highlight differences between ACE and non-ACE students that are not revealed by an analysis of school achievement and participation data alone.

Survey Pilot

Two schools agreed to participate in the high school pilot. Each school was given a protocol (see Appendix J) that contained instructions for survey implementation, questions for students and staff on the survey itself, and a copy of the survey. A total of 78 students completed the survey. Survey results, which are preliminary and for survey testing purposes only, are contained in Appendix K.

Survey Feedback

Feedback from the two sites was relatively limited but included the following suggestions rated to length, clarity of questions, and response categories.

Length

It appears the survey took students between 3 and 10 minutes to complete. This is a short amount of time, given the length of the survey, and suggests that students did not take the survey seriously or may have rushed through it. However, it suggests a need for clear instructions about how to administer the survey to ensure optimal student attention and engagement for the fall data collection.

One of the administrators felt the survey was too long and that the questions seemed repetitive and could be simplified. Student surveys often can seem redundant, but that is by design. In some cases, the questions ask students something similar in several different ways to ensure accuracy of response. The evaluation team reviewed the survey for unnecessary repetition and attempted formatting changes to reduce the number of pages.

Question Content

Students and administrators said the questions were clear and easy to understand. At one school, English language learners had a harder time understanding the questions because of the language barrier, suggesting a need to translate the survey into Spanish for the fall administration.

The open-ended questions appear to be too similar to yield a wide variety of answers. Students often answered the same thing to all three questions (Why did you first choose to come to the program? What is the number one reason you stay in the program? What do you like most about the program?). Because of this difficulty, the evaluation team removed one open-ended question and replaced it with a question about what students liked least about the program.

Response Categories

Some students felt that the response categories were confusing or that it was hard to choose an answer. One administrator suggested that a three-point response scale would be easier and make more sense for students. Typically, a three-point response scale does not show enough variation for the kind of analysis that will be performed on these data. In a final review of the survey, the evaluation team decided to keep the four-point scale for the fall data collection.

In both schools, students asked whether they could choose more than one response category on the question related to where they go after school when they are not at the ACE program. In addition, 19 out of 78 students (or about 25 percent) did not respond to this question. It is unknown whether that is because they do not go anywhere else (e.g., spend five days per week at ACE) or because of being forced to choose one answer. The question was reworded to ask where students *most often* go after school when they are not in the program. This question is important for understanding what ACE students are doing in comparison with non-ACE students after school.

Next Steps

In general, the feedback on the survey was positive and limited in scope. AIR used the feedback to develop a revised version of the survey for use during the 2014–15 evaluation. The pilot, although small, was useful in providing information about the face validity of the survey and giving AIR a sense of which questions were most confusing. With revisions, the team is confident the survey can produce strong data for the large high school study.

Afterward: Cost Study

This section outlines descriptive cost analyses conducted for the 2013–14 SPP grant program funding. SPP funding supported the afterschool activities of a discrete group of students selected for the pilot program; however, each center where SPPs were in place also had traditional Texas ACE programs running as well. The analyses presented in this section reflect only the activities provided for students in the SPP. On the basis of the research team's analysis of 2012–13 SPP grant application budgets and SPP continuation budgets for the 2013–14 year, it was evident that grantees planned to use different strategies to allocate their SPP grant funds across sites and features of their programs. Therefore, analyses were conducted to provide information about how SPP grant funds were used at the site level in 2013–14, with the goal of exploring the following:

- SPP budgets and expenditures by the five major categories reported to TEA: payroll, professional and contract services, supplies and materials, other operating costs, and capital outlay
- SPP budgets and expenditures across program activities: online academic interventions, face-to-face academic interventions, enrichment activities, social services for students and parents, project management, and other
- SPP spending across payroll positions serving SPP (i.e., academic, administrative, and auxiliary)
- Per-pupil expenditures across SPP sites, among students receiving SPP services at an SPP site
- Information related to program sustainability, including the proportion of SPP grant funds allocated to nonrecurring or onetime costs; other sources of funding used to support SPP activities; and whether personnel were paid entirely, or partially, out of SPP budgets

In addition to providing information on SPP site-level spending in 2013–14, these efforts serve as a precursor to the 2014–15 cost analyses, which will explore relationships between site-level expenditures and student outcomes (e.g., regular school-day attendance, STAAR results for mathematics and reading). The goal of those analyses will be to provide insight into an important element of SPP programs—that is, which SPP program elements and features are associated with the most substantial returns in terms of student outcomes. Analyses and data collection efforts for the current year set the stage for 2014–15 cost analyses by first establishing how funds were budgeted and then actually used at the site level.

Data and Methods

SPP budgets contained in 2013–14 grant proposals were reported at the grantee level, not the site level; therefore, it was necessary to collect supplemental financial data about SPP expenditures by site. Researchers collected these data, which were disaggregated by key programmatic functions and payroll functions, and analyzed them

descriptively. The ratio of site-level spending to SPP student enrollment also was assessed to provide a measure of per-pupil spending by site.

To collect site-level spending data systematically, the research team created standardized spreadsheets that requested budget and expenditure data from each grantee and requested that project directors complete separate spreadsheets for each site receiving SPP funding. The spreadsheets were composed of multiple worksheets, with different data collected by worksheet. Specific data elements collected for this report included budget and expenditures by expenditure category (e.g., payroll, professional and contract services) and by program activity (e.g., online academic interventions, project management). Detailed payroll spending and information related to program sustainability (e.g., nonrecurring or onetime costs incurred during the 2013–14 school year, other funding sources for SPP activities) also were requested. Because of complexities involved with tracking and reporting exact amounts of ACE-only funding spent on SPP activities, the total amount of supplemental funds was not requested. Rather, grantees were t instructed to indicate whether specific, ACE-only funding sources were used to support the SPP program.

In April 2014, the evaluation team sent each project director a spreadsheet(s) to be completed for their SPP site(s) and detailed instructions for completing each worksheet. Evaluators requested that the financial information cover all SPP expenditures that occurred between September 1, 2013, and May 20, 2014, allowing researchers to explore how SPP funds were spent during the 2013–14 academic year. Researchers requested that grantees submit this information by June 15, 2014. To facilitate data collection efforts, researchers invited all SPP project directors to join a conference call in mid-April to discuss the data collection spreadsheets and timeline and to address questions or concerns. Project directors, or their designees, from 9 of the 13 grantee organizations participated in the call. The research team also provided technical assistance on completing the spreadsheets by means of e-mails and phone calls with individual project directors.

Most grantees provided the requested information by the deadline, and all grantees submitted information for their sites by the end of June. Data validation procedures illustrated that most of the information provided was correct, with the exception of apparent discrepancies in the reporting of nonrecurring costs (e.g., onetime software costs, iPad or laptop computer expenses). Given concerns about the accuracy of these data, nonrecurring costs are omitted from the discussion of program sustainability.

Use of SPP Funds

Grantee- and Site-Level SPP Budget and Expenditure Data

Table 20 reports average grantee-level spending of approximately $150,000 during the 2013–14 school year, leaving roughly 10 percent of funding for use after May. In other words, on average, grantees spent 90 percent of their SPP funding during the school year.

Table 20. Grantee-Level SPP Budgets and Expenditures (2013–14)

Grantee-Level SPP Budgets and Expenditures	Minimum	Maximum	Mean	Median	Total
Total Reported Budget	$112,110	$217,947	$166,855	$186,819	$2,169,510
Total Spending	$86,284	$198,926	$149,162	$159,940	$1,939,100
Remaining Budget			$17,724	$26,879	$230,409
Percentage of Budget Remaining			11%	14%	11%
N = 13					

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Of the 13 SPP grantees, most reported site-level spending for multiple sites—only two SPP grants were awarded to organizations with only one site. For the 37 SPP sites, Table 21 illustrates an average spending of slightly more than $50,000 per site. The range of spending across sites was considerable, with one site spending less than $10,000 and another spending more than $120,000. Despite this range in spending, mean and median spending were almost equal, implying that there were not many extreme values in terms of site-level spending. It appears that sites' retention of their budget after the school year mirrored the grantee level—that is, on average, sites retained 11 percent of their budgets for use after May 2014, meaning that they spent almost 90 percent of their funding from September to May.

Table 21. Site-Level SPP Budgets and Expenditures (2013–14)

Site-Level SPP Budgets and Expenditures	Minimum	Maximum	Mean	Median	Total
Total Reported Budget	$15,671	$123,051	$58,635	$64,663	$2,169,509
Total Spending	$9,999	$123,051	$52,408	$52,792	$1,939,100
Remaining Budget			$6,227	$11,691	$230,409
Percentage of Budget Remaining after May, 2014			11%	18%	11%
N = 37					

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Table 22 provides further illustration of remaining funds after the academic school year. Most sites (75 percent) retained less than $10,000 of their budget for use after May (of those, two actually went over budget during the 2013–14 school year). Six sites retained between $10,000 and $20,000 (or between 20 percent and 40 percent of their budget), and three sites retained more than 40 percent, or more than $20,000. In general, these data show that, although the majority of sites spent most of their budgeted funds during the school year, roughly one in four sites retained more than $10,000 for after May 2014.

Table 22. Remaining Budget as a Proportion of Total Budget (2013–14)

Remaining Budget Category	Number of Sites	Percentage of Sites[1]
Less than 0% (over budget)	2	5%
20% or less	26	70%
20.1%–40%	6	16%
40.1%–60%	2	5%
60.1%–80%	1	3%
Total	37	100%
N = 37 [1] Total may not equal 100% due to rounding		

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Analysis of SPP Budget Variance

As noted, SPP grantees report their budgets and expenditures across TEA expenditure categories, which include payroll, professional and contracted services, supplies and maintenance, other operating costs, and capital outlay. By collecting these data at the site level, it was possible to go beyond grantee-level spending and compare how sites spent funds across these categories. Table 23 presents the minimum, maximum, mean, median, and total expenditures across sites by expenditure category, and Figure 8 presents the proportion of total spending allocated across expenditure categories.

On average, sites spent the most on payroll, roughly $34,000 per site. The amount spent on professional and contract services was next highest ($11,257), followed by supplies and materials ($3,809). Grantees spent the least amount of money on other expenses ($738), which were expenses that could not be categorized in one of the other categories.

Table 23. Site-Level SPP Expenditures by Expenditure Category (2013–14)

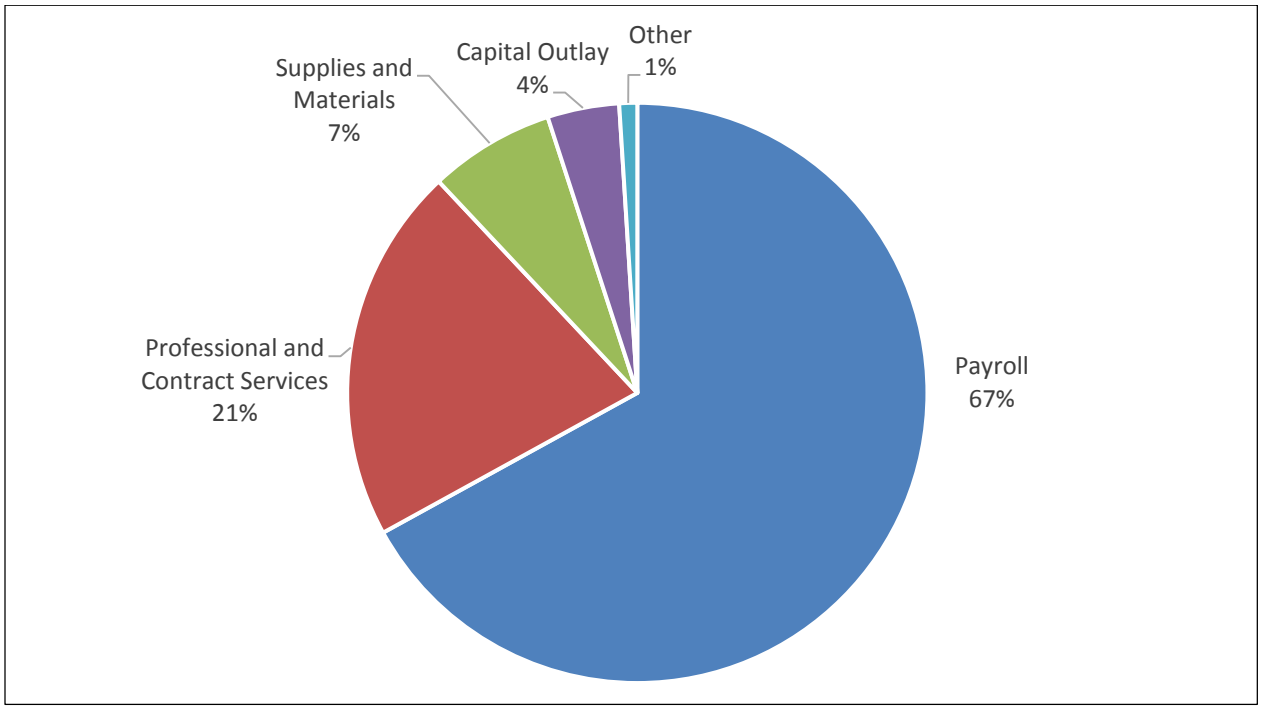
Expenditure Category	Minimum	Maximum	Mean	Median	Total
Payroll	$0	$87,150	$34,329	$32,576	$1,270,183
Professional and Contract Services	$0	$44,237	$11,257	$6,618	$416,514
Supplies and Materials	$0	$31,724	$3,809	$2,282	$140,921
Capital Outlay	$0	$8,033	$2,275	$285	$84,164
Other	$0	$17,280	$738	$0	$27,318
N = 37					

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Considering proportions of spending by expenditure category relative to total spending, sites spent the largest proportion of their grant monies on payroll, which accounted for an average of 67 percent of total spending. This was followed by professional and

contract services (21 percent), and supplies and materials (7 percent).[35] On average, grantees spent the smallest proportion on "other" expenses (1 percent) (see Figure 8).

Figure 8. Average Percentages for Expenditure Categories Across SPP Sites (2013–14)



N = 37

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Table 24 further illustrates spending by expenditure category by providing the number of sites that allocated different proportions of their total spending to a given category—for example, the proportion of overall spending allocated to payroll, professional and contract services, and so on. As with previous results, it is clear from these distributions that the largest proportion of SPP spending went to payroll, though that was not true for all 37 sites. One site allocated less than 20 percent of their SPP expenditures to payroll, and three others spent less than 40 percent. Conversely, expenditures for professional and contract services at one site accounted for more than 80 percent of total spending and for more than 60 percent at three other sites. Spending in the other categories—supplies and materials, capital outlay, and other—accounted for less than 20 percent of most sites' total expenditures, with only three sites spending more than 20 percent on supplies and materials and all 37 sites spending less than 20 percent on capital outlay and other expenses.

[35] Proportions of particular activities or categories were calculated by dividing expenses for a category by overall spending (i.e., expenditures for a category/total reported spending).

Table 24. Ranges of SPP Expenditures by Category, by Number and Percentage of SPP Sites (2013–14)

Expenditure Category	Number of Sites	Percentage of Sites
Payroll		
Less than 20%	1	3%
20.1%–40%	3	8%
40.1%–60%	4	11%
60.1%–80%	22	59%
80.1%–100%	7	19%
Professional and Contract Services		
Less than 20%	23	62%
20.1%–40%	9	24%
40.1%–60%	1	3%
60.1%–80%	3	8%
80.1%–100%	1	3%
Supplies and Materials		
Less than 20%	34	92%
20.1%–40%	3	8%
Capital Outlay: Less than 20%	37	100%
Other: Less than 20%	37	100%
N = 37		

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Across Activity Type

On average, sites spent the most on face-to-face interventions, at roughly $25,000 per site. The amount spent on project management was next highest ($10,600), followed by enrichment activities and other spending ($6,056 and $5,020, respectively). Grantees spent the least amount of money on social services for students and families ($674).

In terms of proportions of overall spending, grantees spent the largest proportion of their grant monies on face-to-face interventions, allocating, on average, 48 percent of their total spending to face-to-face activities (Table 25). This was not true of all sites; spending on face-to-face interventions accounted for almost all of some sites' spending, whereas others allocated very little or none of their total spending to face-to-face activities. For online interventions—the other instructionally related activity—sites allocated about 9 percent of their total expenditures. As with face-to-face interventions, spending varied across sites. For example, some sites spent no money on online activities, and at least one site spent almost 40 percent on online interventions. Among noninstructionally related activities, project management constituted 20 percent, enrichment constituted 12 percent, and other activities constituted 10 percent of total

spending on average. Spending for social services for students and families made up the smallest proportion of overall spending (1 percent). Generally, sites spent 60 percent of their SPP funding on instructionally related activities (i.e., face-to-face and online interventions), with the remaining 40 percent split across noninstructional activities.

Table 25. Proportion of SPP Expenditures by Activity Type (2013–14)

Activity Type	Minimum	Maximum	Mean	Median
Enrichment Activities	0%	48%	12%	4%
Face-to-Face Interventions	0%	96%	48%	39%
Online Interventions	0%	38%	9%	0%
Project Management	0%	60%	20%	9%
Social Services for Students and Families	0%	21%	1%	0%
Other	0%	37%	10%	7%
N = 37				

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Table 26 shows the proportion of total spending by activity type. Five sites allocated between 80 percent and 100 percent of their spending to face-to-face activities, and six sites allocated less than 20 percent. Conversely, expenditures for online interventions accounted for between 20 percent and 40 percent of total spending at 10 sites. Spending in noninstructional categories—enrichment, project management, social services for students and families, and other—accounted for no more than 60 percent of any particular site's total expenditures. Three sites spent between 40 percent and 60 percent on project management, and one site spent the same proportion on enrichment activities. With regard to social services and other, most sites allocated less than 20 percent of their total expenditures to those activities, although one site reported spending more than 20 percent on social services, and six sites spent more than 20 percent on other expenses. As with the instructional categories, the remaining spending categories varied. Three sites spent between 40 percent and 60 percent of their funding on project management, and one site spent between 20 percent and 40 percent of their funding on social services for students and parents.

Table 26. SPP Expenditures by Activity Type, by Number and Percentage of SPP Sites (2013–14)

Activity Type	Number of Sites	Percentage of Sites[1]
Enrichment Activities		
Less than 20%	25	68%
20.1%—40%	11	30%
40.1%—60%	1	3%
Face-to-Face Interventions		
Less than 20%	6	16%
20.1%—40%	10	27%
40.1%—60%	6	16%
60.1%—80%	10	27%
80.1%—100%	5	14%
Online Interventions		
Less than 20%	27	73%
20.1%—40%	10	27%
Project Management		
Less than 20%	24	65%
20.1%—40%	10	27%
40.1%—60%	3	8%
Social Services for Students and Families		
Less than 20%	36	97%
20.1%—40%	1	3%
Other		
Less than 20%	31	84%
20.1%—40%	6	16%

N = 37
[1] Total may exceed 100 percent due to rounding.

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Across Payroll Categories

This section describes results from analyses of spending by payroll category, with most analyses using aggregated payroll categories (i.e., academic, administrative, and auxiliary[36]). Almost 70 percent of SPP funds were allocated to academic personnel, with most of the remaining payroll funding allocated to administrative personnel (28 percent) and a small percentage allocated to auxiliary staff (3 percent) (see Table 27).

[36] Auxiliary services refer to nonacademic support services provided to students in the SPP, including social services (e.g., counseling, home visits), nursing services, and similar services.

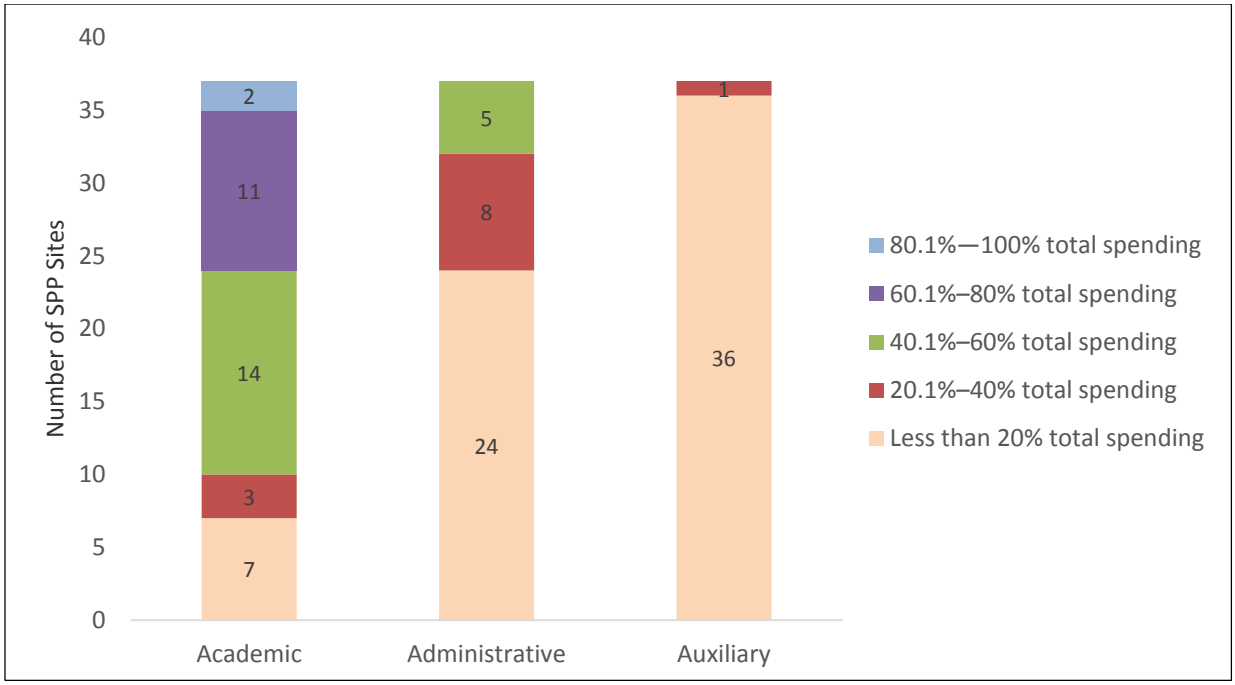
Table 27. Proportion of Expenditures by Payroll Category (2013–14)

Payroll Category	Minimum	Maximum	Mean	Median
Academic	19%	100%	68%	76%
Administrative	0%	78%	28%	18%
Auxiliary	0%	28%	3%	0%
N = 37				

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

As Figure 9 demonstrates, spending for academic positions varied, with seven sites allocating less than 20 percent of their total spending and two sites allocating more than 80 percent of their payroll spending to academic personnel. In addition, despite the tendency for most sites to use the majority of their payroll funds on academic personnel, five sites spent more than 40 percent of their funding on administrative personnel.

Figure 9. Expenditures by Payroll Category, by Number and Percentage of SPP Sites (2013–14)



N = 37

SOURCE: SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Of all payroll positions, the largest proportion of spending went to teachers (42 percent of payroll expenditures). The remaining academic positions (i.e., college and career readiness specialists, educational aids, and tutors) received 13 percent of payroll spending. In the administrative category, almost a quarter (23 percent) of payroll spending was split between project directors and site coordinators. The remaining

payroll funding was allocated to family engagement specialists (3 percent), secretaries and administrative assistants (2 percent), data entry clerks (1 percent), grant accountants and bookkeepers (1 percent), and evaluation specialists (3 percent). A smaller proportion of SPP funding was allocated to auxiliary positions (e.g., school nurses, bus drivers, and social workers), who, combined, accounted for roughly 3 percent of payroll spending.

Per-Pupil Spending

These analyses sought to address questions about variability in spending per student across sites, including the following:

1. Did SPP sites vary in per-pupil spending?
2. Did sites vary in terms of differences between actual per-pupil spending and expected per-pupil spending based on original budgets?

The following cost-per-student ratio was calculated as site-level spending relative to the number of SPP students. The research team also calculated ratios between SPP enrollment and total site-level budgeted amounts to obtain per-pupil budgets (Table 28).

Table 28. Per-Pupil Spending, Budgeted, and Differences Between Spending and Budgeted (2013–14)

	Minimum	Maximum	Mean	Median	Total
SPP Students Only (*N* = 4,413)					
Per-Pupil Spending	$118	$4,061	$651	$401	$24,105
Per-Pupil Budgeted	$121	$4,061	$710	$471	$26,254
N = 37					

SOURCE: SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Analyses of student attendance and site-level expenditures demonstrate that per-pupil spending varied substantially across sites, with site-level spending ranging from $118 to $4,061 per SPP student, indicating that the cost of providing SPP interventions was substantially higher per student at some SPP sites than others. Almost 50 percent of sites (*n* = 18) spent $400 or less per SPP student, and almost 25 percent of sites (*n* = 9) spent more than $600 per SPP student (Table 29).

Table 29 further illustrates per-pupil spending by providing the number of sites that fell within specific per-pupil spending ranges (e.g., $100 or below, between $101 and $200, and so on), along with average site-level SPP student enrollment within those spending categories. The results suggest an inverse relationship between per-pupil costs and student enrollment, meaning that programs with fewer SPP students spent more per pupil, on average, than did sites with larger SPP student populations. Considering per-pupil spending relative to site-level SPP budgets, roughly 40 percent of sites budgeted $400 or less per student and 35 percent of sites budgeted more than $600 per SPP student (Table 29).

Table 29. Per Pupil SPP Spending, Budgeted, and Differences by Number and Percentage of Sites (2013–14)

	SPP Students Only		
	Number of Sites	Percentage of Sites[1]	Average Student Enrollment[2]
Per-Pupil Expenditures			
$100 or below	0	0%	-
$101–$200	3	8%	298
$201–$300	11	30%	124
$301–$400	4	11%	131
$401–$500	7	19%	118
$501–$600	3	8%	76
More than $600	9	24%	64
Per-Pupil Budgets			
$100 or below	0	0%	-
$101–$200	1	3%	552
$201–$300	5	14%	133
$301–$400	8	22%	133
$401–$500	9	19%	136
$501–$600	3	8%	112
More than $600	11	35%	65

[1] Total may exceed 100 percent due to rounding.

[2] Average student enrollment figures reflect the average number of students served by sites within the various per-pupil spending, per-pupil budgets, and budgets-versus-spending categories.

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Program Sustainability

In addition to collecting information on site-level spending by activity type and payroll category, the supplemental financial data collection form sent to all grantees requested data on nonrecurring costs incurred in 2013–14. This collection effort attempted to capture information on the proportion of SPP funds that could be considered onetime or infrequent expenses (e.g., building or supplies costs), in contrast to costs that would be expected to reoccur annually (e.g., payroll costs). By collecting these data, the research team hoped to assess the total amount of funding that would be required to sustain program activities in future years, by subtracting nonrecurring costs from total costs to get a sense of what would be required in future years to sustain SPP program continuation.

Although all but five sites reported nonrecurring costs for 2013–14, one site reported total nonrecurring costs in excess of their total spending, and 12 sites (32 percent) indicated nonrecurring spending on supplies that accounted for more than 100 percent of their reported spending on supplies. In addition, narrative descriptions of the

nonrecurring costs demonstrated that, although sites reported on onetime or infrequent costs such as “tablets and white boards,” they also reported costs that will likely reoccur, such as professional development opportunities for faculty and staff. Because of these issues related to the reporting of nonrecurring costs, it was not possible to obtain a valid estimate of the proportion of nonrecurring to total costs in 2013–14.

Although nonrecurring costs were not available for 2013–14, two other sources of information were used to address program sustainability: (1) information on whether sites used ACE-only funding sources (e.g., federal, state, local, or private funds) to support SPP activities in 2013–14 and (2) information on whether staff were paid completely or partially with SPP funds.

With regard to other funding sources, all sites reported using 21st CCLC funding to support SPP activities in 2013–14. Aside from 21st CCLC funding, the majority of sites (70 percent, *n* = 26) reported using at least one ACE-only/non-21st CCLC funding source to support SPP activities (Table 30).

Table 30. Number and Proportion of Sites Using Other Funds to Support SPP (2013–14)

Funding Sources	Number of Sites	Percentage of Sites[1]
Federal Funds		
None	20	54%
1	12	32%
2	1	3%
3	4	11%
State Funds		
None	25	68%
1	11	30%
2	1	3%
Local Funds		
None	35	95%
1	2	5%
Total Other Funds (non-21st CCLC)		
None	11	30%
1 to 3 funding sources	21	57%
4 to 6 funding sources	5	13%

N = 37
[1] Totals may exceed 100 percent due to rounding.

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Almost half reported using at least one federal funding source, with Title I Program funds (46 percent, *n* = 17 sites) being the most commonly reported. A third of the SPP sites reported using other state-level funds, including Compensatory Education Funds

(19 percent, *n* = 7 sites) and General Funds (11 percent, *n* = 4 sites). Only two sites (5 percent) reported using local funds (Table 31).

Table 31. Number and Proportion of Other Funds Used to Support SPP (2013–14)

Funding Sources	Number of Sites	Percentage of Sites
Federal Funds[1]		
Gear Up	4	11%
Bilingual Education Program Funds	5	14%
Migrant Education Program Funds	5	14%
Title I Program Funds	17	46%
State Funds[1]		
Foundation School Program Funds	1	3%
Public Charter Budget	1	3%
General Funds	4	11%
Compensatory Education Funds	7	19%
Local Funds	2	5%
21st CCLC	37	100%
[1] Grantees may use more than one federal or state funding source to support SPP activities. *N* = 37		

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

An additional source of information on program sustainability was derived from reporting on different payroll positions—specifically, the extent to which certain personnel were paid entirely or partially with SPP grant funding. This information is important because positions that are not totally dependent on grant funding may be more sustainable going forward than are 100 percent grant-funded positions. Most sites reported academic positions as 100 percent grant funded (81 percent, *n* = 30), and spending for administrative and auxiliary positions was more evenly split between being completely and only partially grant funded (Table 32).

Table 32. Proportion of Payroll Categories That Were 100 Percent Versus Less Than 100 Percent SPP Funded (2013–14)

Payroll Category	100% Grant Funded	Less Than 100% Grant Funded
Academic (*N* = 37)	81%	19%
Administrative (*N* = 31)	42%	58%
Auxiliary (*N* = 7)	43%	57%
N = 37		

SOURCE: 21st CCLC evaluation spreadsheets completed by SPP project directors, 2013–14.

Key Findings: Cost Analysis of SPP Programs 2013–14

Most sites expended grant money closely in line with their proposed budget, using almost 90 percent of their allocated funds from September 2013 to May 2014. Grantees tended to spend the largest proportion of grant money on payroll, followed by professional and contract services. On average, grantees spent roughly 60 percent of their grant budgets on academic activities, with almost 50 percent allocated specifically to face-to-face interventions.

Per-Pupil Spending

Data on student attendance and program expenditures indicate that per-pupil costs varied substantially by site, with site-level expenditures ranging from $118 to $4,061. Sites with lower SPP student enrollment incurred generally higher per-student costs. When compared with costs that would have been expected per pupil on the basis of SPP budgets, most grantees' budgets allowed for higher per-student spending than what was actually spent from September 2013 to May 2014.

Program Sustainability

Although the research team was not able to estimate nonrecurring costs, other sources of information provide encouraging data related to program sustainability. For one, information on sites' use of different funding sources to support SPP activities implies that most sites diversify their financial support for SPP activities. In addition, although most academic positions are funded completely by using SPP funding, more than half of administrative and auxiliary positions were funded at less than 100 percent SPP funding.

References

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *British Journal of Educational Psychology*, *72*, 261–278.

Auger, A., Pierce, K. M., & Vandell, D. L. (2013, April). *Participation in out-of-school settings and student academic and behavioral outcomes.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Birmingham, J., Pechman, E. M., Russell, C. A., & Mielke, M. (2005). *Shared features of high-performing afterschool programs: A follow-up to the TASC evaluation.* Washington, DC: Policy Studies Associates.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences.* Mahwah, NJ: Erlbaum.

California Department of Education. (2002). *Evaluation of California's after school learning and safe neighborhoods partnerships program: 1999–2001*. Sacramento, CA: California Department of Education. Retrieved from www.cde.ca.gov/ls/ba/as/documents/execsummary.asp.

Center for Public Policy Priorities. (2013). *Investing in our future: 2013 State of Texas Children.* Austin, TX: Author.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper Collins.

Csikszentmihalyi, M., & Schneider, B. (2000). *Becoming adult: How teenagers prepare for the world of work*. New York, NY: Basic Books.

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

The Equity and Excellence Commission. (2013). *For each and every child—A strategy for education equity and excellence*. Washington, DC: Author.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., et al. (2012). *Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review*. Chicago, IL: University of Chicago Consortium on Chicago School Research.

Glisson, C. (2007). Assessing and changing organizational culture and climate for effective services. *Research on Social Work Practice, 17*(6), 736–747.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.

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.

Howard, G. S. (1980). Response shift bias- -a problem in evaluating interventions with pre/post self-reports. Evaluation Review 4(1): 93 - 106.

Kataoka, S., & Vandell, D. L. (2013). Quality of afterschool activities and relative change in adolescent functioning over two years. *Applied Developmental Science, 17*(3), 123–134.

Kauh, T. J. (2011). *AfterZone: Outcomes for youth participating in Providence's citywide afterschool system*. Philadelphia, PA: Public Private/Ventures.

Larson, R., & Hansen, D. (2005). The development of strategic thinking: Learning to impact human systems in a youth activism program. *Human Development, 48*(6), 327–349.

Linacre, J. M. (2005) *WINSTEPS* [Computer program]. Chicago: Winsteps.com.

Linacre, J. M., & Wright, B. D. (2004). Construction of measures from Many-Facet data. In E. V. Smith, Jr., & R. M. Smith (Eds.), *Introduction to Rasch measurement* (pp. 296–321). Maple Grove, MN: JAM Press.

Miller, B. (2003). *Critical hours: Afterschool programs and educational success.* Boston, MA: Nellie Mae Education Foundation.

Naftzger, N., Manzeske, D., Nistler, M., & Swanlund, A. (2012). *Texas 21st Century Community Learning Center: Interim evaluation report.* Chicago, IL: American Institutes for Research. Retrieved from http://tea.texas.gov/Reports_and_Data/Program_Evaluations/Out-of-School_Learning_Opportunities/Program_Evaluation__Out-of-School_Learning_Opportunities/

Naftzger, N., Manzeske, D., Nistler, M., & Swanlund, A. (2013). *Texas 21st Century Community Learning Center: Year 2 evaluation report.* Chicago, IL: American Institutes for Research. Retrieved from http://tea.texas.gov/Reports_and_Data/Program_Evaluations/Out-of-School_Learning_Opportunities/Program_Evaluation__Out-of-School_Learning_Opportunities/

National Research Council. (2002). *Community programs to promote youth development*. Washington, DC: National Academies Press.

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist, 46*(3), 168–184.

Schafer, J. L., & Kang, J. D. (2008). Average causal effects from nonrandomized studies: A practical guide and simulated case study. *Psychological Methods 13(*4), 279–313.

Shernoff, D. J., & Vandell, D. L. (2007). Engagement in afterschool program activities: Quality of experience from the perspective of students. *Journal of Youth Adolescence*, *36*, 891–903.

Smith, C. (2007, March). *Predictors of quality at the point-of-service provision: Empirical and methodological background for the YPQA field trial*. Presented at the biennial meeting of the Society for Research in Child Development, Boston, MA.

Smith, C., Akiva, T., Sugar, S., Lo, Y.-J., Frank, K. A., Peck, S. C., et al. (2012). *Continuous quality improvement in afterschool settings: Impact findings from the Youth Program Quality Intervention study*. Washington, DC: Forum for Youth Investment.

Texas Education Agency. (2014). *Enrollment in Texas Public Schools 2012–13*. Retrieved from http://www.tea.state.tx.us/acctres/Enroll_2012-13.pdf

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

Appendices

Appendix A. ACE Program's Critical Success Model

Texas ACE Critical Success Factors

Outcomes	Critical Success Factors (behaviors)	Critical Success Factors Performance Indicators	Milestones (grantees)	Milestone Performance Indicators (measure)
• Improve Academic Performance • Improve Attendance • Improve Behavior • Increase Promotion Rates • Increase Graduation Rates	**Critical Success Factor #1: Student and Family Engagement**			
	• Students and families actively participating and engaged in learning • Students and families displaying leadership roles, volunteering to participate and lead activities	• Increased student and family attendance in afterschool programs • Students mentoring other students • Students and families facilitating activities Measurement Tool • Instructor surveys/self assessment • Principal/Project Director survey • Observation/on-site visit	• Utilize innovative instructional techniques for academic and enrichment activities based on research and best practices	• Activity Tracking - TX21st (Three times per year- Summer, Fall, & Spring) • Curriculum/Lesson Plans
	Critical Success Factor #2: School Involvement			
	• Students increased sense of involvement in school	• Number of students participating in extracurricular activities • Increased number of mentors Measurement Tool • Student/Family surveys • Teacher surveys	• Provide adult advocates, based on student need and in accordance with best practices	• Number of meetings with students • Number of contacts made with Families, teachers, school day staff
	Critical Success Factor #3: Assessment Data			
	• Use of assessment data to revise/ reevaluate student services	• Changes in student activities following re-assessment Measurement Tool • Document analysis of program files • Observation/on-site visits	• Conduct ongoing/continuous assessment to determine need and improve targeted services	• Methods of assessment: pre/ post tests, needs assessments, case plans, etc. • Use of PRIME Assessment
	Critical Success Factor #4: Professional Development Impact			
	• Implementation of strategies learned through training • Noticeable difference in educational instruction (teaching methods)	• Changes in methods of instruction based on training Measurement Tool • Self assessments • Supervisor assessments	• Provide all required training opportunities for staff development	• Number of trainings • Schedule of trainings • Staff sign in sheets • Participant surveys • MyTexasACE Training Report

Appendix B. Study Methods

The evaluation methods for 2012–13 and 2013–14 were similar and included a combination of analysis of Texas administrative data, student and activity leader surveys, and site visits. The methods are presented in detail in this appendix along with information about differences between the two years of the study.

Data Collection Methods

Texas Administrative Data. A substantial amount of information housed in TEA's administrative data systems was obtained to support the evaluation. These data include the following:

TX21st. TX21st is a Web-based data collection system developed and maintained by TEA to report required data into the federal 21st CCLC Profile and Performance Information Collection System database. Throughout the program year, TX21st collects data directly from grantees on a broad array of ACE and SPP program characteristics, student demographics, and programs and activities. Data extracted from the tracking system were used to explore levels of enrollment and attendance in SPP-funded activities and how the attendance profiles of SPP and ACE participants differed.

Additional TEA Data. The participant and impact analyses described in this report included variables on student demographics, discipline incidents, school-day attendance, and grade promotion from the Public Education Information Management System (PEIMS). Student achievement data on annual assessments also were used to support the matching of SPP and ACE youth for the impact analyses. School-level performance data from the Texas Academic Performance Report[37] also were used to support matching efforts in 2012–13.

Project Director Interviews. Project directors for the 15 grantees (13 in 2013–14) that received SPP funding were interviewed over the telephone in January 2013 in order to learn about the SPP programs they planned to implement. (Prior to these interviews, the grant applications from the 15 grantees were reviewed.) They were conducted again in December 2013 as part of the 2013–14 evaluation to understand any differences between the two years of program implementation. The interviews were helpful in understanding how the SPP programming was designed and in selecting appropriate SPP programs for site visits planned for spring 2013 and 2014.

Recruitment Survey. A brief survey was administered to the grantee project directors in December 2012 and again in October 2013 to learn how they were identifying students for participation in the program. Responses were received from 13 of the 15 grantees that had received SPP funding in 2012–13 and from all 13 grantees that had received SPP funding in 2013–14. These data were used to understand better who was

[37] Prior to 2012–13, the Texas Academic Performance Report was known as the *Academic Excellence Indicator System*.

being targeted for program participation as well as to inform the ACE-only comparison group for student surveys described in more detail in the following section.

Youth Mindsets and Behaviors Survey and Related Activity Leader Reports. A consistent goal of the evaluation has been to explore further the SPP theory of change, showing that high levels of engagement in high-quality afterschool programs can contribute to changes in academic mindsets and behaviors that then lead to generalized school success. As a preliminary step in that process, the evaluation team endeavored to develop new measures, based on existing valid and reliable surveys, to assess student functioning associated with academic mindsets and behaviors. The survey included items to measure the following constructs: *Effort and Persistence*, *Learner Behaviors*, *Self-Efficacy/Self-Competence*, and *Mastery Orientation*.

Two approaches for collecting these data were used. For students in Grades K–3, who were too young to take a survey reliably, a report was developed for SPP activity leaders to complete about their Grade K–3 students. The report included questions on student functioning related to two of the survey constructs: *Effort and Persistence* (for which there were seven survey items) and *Learner Behaviors* (for which there were 10 items). Students in Grades 4–12 completed the Mindsets and Behaviors Survey themselves, online or on paper. The survey for these students included items on all of the constructs.

During 2012–13, the activity reports and the student survey were administered first in February 2013, as a pretest, and again in May 2013, as a posttest. This time frame was a result of the SPP programming schedule: With programming starting in early 2013, the pre- and posttests were administered in relatively close proximity. The pretest survey and activity reports were collected at the time programming began and at the time programming was concluding for the school year. Because of the short assessment window, 2012–13 findings related to pre- and posttest differences should be interpreted with caution. In 2013–14, the two types of surveys were administered first in November 2013 and again in April and May 2014.

Both activity leader reports and youth surveys were administered at all centers with SPP programming, though only to a sample of SPP and ACE-only program participants. Using a stratified, randomized sampling approach, based on data housed in TX21st, AIR selected 50 SPP participants to take the survey. The evaluation team calculated the proportion of students in each grade and for each gender that matched the overall distribution of SPP students. A sampling of those students then was selected. A sample of alternate SPP students also was identified in the event that a student from the original sample either was not enrolled or had left the SPP program prior to data being collected.

In addition to developing a sample of SPP student survey respondents, the evaluation team asked ACE center staff to identify up to 30 ACE students who were not enrolled in the SPP programming. The evaluation team also provided numbers of students and criteria for the 30 non-SPP students whom sites could select to take the survey (Grades 4–12) or have an activity leader or teacher fill out a survey on their behalf (Grades K–3).

The evaluation team asked that all sites survey 30 ACE-only students, even if the site had fewer than the desired 50 SPP students. The evaluation team provided the distribution of students in each grade and gender that matched the distribution of the SPP students and that added up to 30 total ACE-only students. In addition to the distribution of students, the evaluation team also included academic criteria for selecting the ACE-only students. To determine these criteria, the evaluation team surveyed project directors to determine how they identify students for the SPP program. Each site received a letter that included selection criteria along with the numerical distribution of students to guide the selection of ACE-only students to take the survey.

In 2012–13, surveys and activity leader reports for ACE-only students were completed in December 2012, two months before SPP surveys and activity leader reports were completed. Both pre- and posttest activity leader report data were collected on 821 Grade K–3 students and 1,175 Grade 4–12 students. These data were used to assess preliminary SPP program impact during the spring 2013 implementation period.[38]

In 2013–14, surveys were administered at the same time for all students. Both pre- and posttest activity leader report data were collected for 1,952 SPP and ACE-only students. Survey data were used to assess preliminary SPP program impact during the 2013–14 school year.

Activity Leader Surveys. Activity leader surveys were administered at all ACE centers with SPP funding to *all* SPP activity leaders at the center and to the activity leaders of ACE-only activities that were offered at the same center if that activity leader had a connection to a student who took a student survey. Each center was given the list of ACE-only students who took the student survey and was asked to identify at least one ACE-only activity leader associated with each student to complete the activity leader survey. This was to ensure that the activity leader responses and student responses would be more likely to refer to the same programs and activities. The survey asked about activity leader and program characteristics and also about indicators of program quality. Eight constructs related to quality were included in the survey, each consisting of six to nine items. Three constructs were related to organizational quality: *Internal Communication*, *Program Climate*, and *Program Challenges*. Five constructs were related to activity quality: *Intentional and Responsive Instruction*, *Individualized and Contingent Instruction*, *Access to Student Data*, *Academic Learning Activities*, and *Youth Development Activities*.

In 2012–13, 443 surveys were completed. The survey respondents included 246 SPP activity leaders and 197 ACE-only activity leaders. In 2013–14, 392 surveys were completed. The survey respondents included 180 SPP activity leaders and 182 ACE-only activity leaders. Thirty activity leaders classified themselves as both SPP and ACE-only activity leaders.

[38] As noted earlier, there was a very short time frame between pretest and posttest administrations because of the SPP programs not being implemented until early in 2013.

Site Visits. During both evaluation years, site visits were conducted in the spring (2013 and 2014, respectively) at one ACE center with SPP funding, representing each of the SPP grantees. One site visit per grantee agency was conducted (e.g., 15 sites visits in 2012–13 and 13 sites visits in 2013–14). In addition, in 2013–14, the evaluation team identified a matched sample of ACE-only centers within the same grantee when possible for site visits, so in 2013–14, the team conducted 26 site visits to 13 SPP centers and 13 ACE-only centers. Project director interviews were used to assess the extent to which SPP activities were being offered in a uniform manner across participating sites and to provide recommendations for sites with strong management and effective implementation of the SPP program. Site selection was based on these recommendations and the desire to visit sites serving a range of elementary, middle school, and high school students.

During 2012–13, the site visits were one day. In 2013–14, the evaluation team expanded the visits to two days to accommodate more observations and more interviews and to add the opportunity to bring in a second rater. Site visit activities included interviews with activity leaders and site coordinators, observations of program sessions, and administration of a student survey to Grade 4–12 students who participated in the sessions that were observed.

Observations. In 2013–14, one or two observations of SPP activity sessions were conducted at each site for a total of 22 observations. In 2013–14, because of the longer site visits, the evaluation teams were able to observe more programs (typically three to five) for a total of 89 observations.

Observation Measures. As noted in the body of the report, observations were guided by three instruments.

- **PQA.** The primary observation tool employed was the PQA,[39] an instrument developed by the High Scope Education Foundation and now administered by the David P. Weikart Center for Youth Program Quality (http://cypq.org/downloadpqa). The tool frequently is used to assess the quality of extended learning programs at the POS.
- **APT-O.** This tool, specifically the section on academic skill building, was used to measure the types of support and opportunities afterschool staff can provide to support skill development in particular content areas. The APT-O (http://www.niost.org/apt) is a comprehensive observation tool developed by the National Institute on Out-of-School Time for the Massachusetts Department of Elementary and Secondary Education. The tool was designed to support state efforts to improve 21st CCLC. The APT-O was selected to supplement the PQA because it includes measures of skill building in reading and mathematics.
- **OCES.** The OCES (developed by the University of Virginia Social Development Lab, http://www.socialdevelopmentlab.org/resources/measures/oces/), an

[39] As noted earlier, two versions of the PQA were used to support observations conducted at SPP sites. The School-Age PQA was used in centers serving elementary students. The Youth PQA was employed in centers serving middle and high school students.

adaptation of the National Institute of Child Health and Human Development Early Child Care Research Network Classroom Observation Scale, was used to measure the extent to which youth participating in a given activity were engaged in program activities, attentive, self-reliant in performing activity tasks, and compliant with requests and directives from activity leaders. During the observations, observers applied ratings for each of the constructs on the three measures as well as maintaining a narrative description of the observed, describing activities, materials and resources, teacher-and-student interactions, student interactions, and student engagement. The measure consists of five items: engagement, attention, self-reliance, compliance, and disruptive behavior. Each item was rated on a four-point Likert-type scale.

The activities observed were primarily either academic enrichment (73 percent) characterized by an intentional effort to build youth skills in a specific academic content area or nonacademic enrichment (26 percent), which were more apt to emphasize the provision of developmentally appropriate activities that more generally supported youth development. One small-group tutoring session focusing on mathematics also was observed. The mean number of youth participating in an observed activity was 10 (ranging from 1 to 60), and the median number of staff was one (ranging from one to five). Differences in terms of activity type and the number of staff and youth participating in observed sessions between SPP and ACE-only centers are outlined in Table B-1a and B-1b. As shown, activities observed in SPP programs were more likely to be classified as academic enrichment ($p < 0.05$, chi-square = 8.851, df = 2) and have more staff present on average ($p < 0.001$, $t = 4.498$, df = 55.646) than were activities observed in ACE-only centers.

Table B-1a. Activity Type Associated With Observed Activities by Center Type (2013–14)

Activity Type	Activities in SPP Centers (*n* = 45)		Activities in ACE-only Centers (*n* = 44)	
	Number	Percentage	Number	Percentage
Academic enrichment	39	86.7%	26	59.1%
Nonacademic enrichment	6	13.3%	17	38.6%
Tutoring	0	0.0%	1	2.3%

SOURCE: TEA, TX21st, 2013–14.

Table B-1b. Activity Participants Associated With Observed Activities by Center Type (2013–14)

Youth and Staff Participants	Activities in SPP Centers (*n* = 45)	Activities in ACE-only Centers (*n* = 44)
Mean number of youth	10.8	10.1
Mean number of staff	1.6	1.0

SOURCE: TEA, TX21st, 2013–14.

Youth Surveys. Observations of activities serving students in Grades 4 and above included an end-of-session student engagement survey. The survey is similar to one employed by Shernoff and Vandell (2007) and includes eight questions on students' concentration, enjoyment, and interest during the session (see Appendix C for the specific survey items). The survey assesses a cognitive, as opposed to a behavioral, definition of engagement and relies on self-report rather than an observer scanning for and recording engagement levels of participating youth. In total, 86 student surveys were collected for nine SPP activities provided at nine centers in 2012–13. On average, nearly 10 student surveys were collected at each of the nine centers. In 2013–14, student surveys were much more extensive, being administered to 401 students. The higher number was because of the higher number of observations.

Interviews. Interviews with site coordinators and observed activity leaders were conducted at each SPP site visit. The site coordinator interviews addressed program design, staffing practices, implementation strategies, and issues that might impact programming. Interviews with activity leaders of the observed sessions were conducted to learn more about the purpose of the sessions and what the activity leader perceived to be successful or unsuccessful components of the session.

Appendix C. Staff Survey Responses by Item and Staff Type, 2013–14

Table C-1. *Intentional and Responsive Instruction*, 2013–14

How often do you lead or participate in OST program activities that are…	SPP Staff				ACE-Only Staff			
	Never Occurs	Occurs Occasionally	Occurs Regularly	Always Occurs	Never Occurs	Occurs Occasionally	Occurs Regularly	Always Occurs
1. Based on written lessons outlining the specific purposes of a given activity session, assignment, or project	1.4%	9.9%	31.8%	56.7%	1.7%	13.6%	37.2%	47.3%
2. Designed to support specific, measurable learning goals for students	0.0%	4.3%	25.0%	70.6%	1.7%	7.0%	35.6%	55.5%
3. Planned in advance	0.0%	3.4%	23.7%	72.7%	2.2%	7.3%	28.6%	61.7%
4. Meant to extend and enhance skills developed in prior activity sessions or lessons	0.0%	4.8%	28.1%	66.9%	2.8%	6.7%	29.3%	61.0%
5. Explicitly meant to promote skill building and mastery in relation to one or more state or district standard(s)	0.0%	2.4%	24.0%	73.5%	4.1%	11.1%	25.8%	58.8%
6. Discussed and planned with other ACE/SPP staff prior to activity session delivery	3.4%	21.7%	34.1%	40.5%	7.1%	25.1%	27.5%	40.1%
7. Discussed and planned with other school-day staff prior to activity session delivery	12.8%	28.2%	28.2%	30.6%	17.4%	29.6%	24.5%	28.3%
8. Informed by levels of student engagement in previous activity sessions covering similar content	0.4%	11.2%	44.1%	44.1%	5.8%	14.0%	36.2%	43.8%
9. Intentionally designed to address prior student mistakes or misunderstandings	1.4%	12.1%	36.4%	50.0%	4.1%	19.5%	34.9%	41.4%

Table C-2. *Individualized and Contingent Instruction*, 2013–14

Please indicate the degree to which you know the following things about the students you work with when providing OST programming:	SPP Staff				ACE-Only Staff			
	I know this for None of my students.	I know this for Some of my students.	I know this for Most of my students.	I know this for All of my students.	I know this for None of my students.	I know this for Some of my students.	I know this for Most of my students.	I know this for All of my students.
1. I usually know if a student is struggling with academic content during the regular school day.	3.3%	23.0%	32.6%	40.8%	2.6%	21.4%	46.9%	28.8%
2. I know how a student performed on state- and district-level assessments.	19.5%	18.0%	18.5%	44.0%	26.0%	28.8%	22.5%	22.5%
3. I am generally aware of the specific academic concepts and skills a student is struggling with.	2.8%	16.2%	35.4%	45.4%	5.2%	19.6%	45.7%	29.4%
4. I understand the learning styles of each student I work with.	1.9%	18.6%	44.4%	34.9%	3.1%	22.3%	41.6%	32.9%
5. I am aware of the strengths of each student.	0.9%	17.0%	35.0%	46.9%	3.0%	13.8%	43.3%	39.7%
6. I recognize areas a student needs to improve in.	0.0%	10.4%	32.2%	57.3%	3.0%	11.0%	44.7%	41.1%
7. I have a good understanding of how a student feels about school.	0.0%	11.9%	34.2%	53.8%	1.1%	9.4%	45.2%	44.1%
8. I know how a student's prior experiences with academic content/concepts should impact individual instruction with that student.	3.8%	19.6%	30.6%	45.9%	8.4%	24.0%	37.0%	30.5%

Table C-3. *Access to Student Data*, 2013–14

Please indicate whether you receive each of the following, and to what extent you use it in planning for the activities you provide:	SPP Staff			ACE-Only Staff		
	Do Not Receive	Occasionally Use	Often Use	Do Not Receive	Occasionally Use	Often Use
1. Individual student academic plans	36.4%	20.0%	43.5%	49.2%	26.4%	24.2%
2. Students' state assessment scores	34.0%	20.7%	45.2%	58.7%	19.0%	22.1%
3. Students' scores on state- or district-level assessments	32.9%	16.7%	50.2%	57.5%	19.6%	22.7%
4. Students' grades	31.7%	21.5%	46.6%	37.0%	25.8%	37.0%
5. Input from students' school-day teachers	21.2%	31.6%	47.0%	26.6%	34.0%	39.3%

Table C-4. Service Delivery Practices, 2013–14

In the activities <u>you</u> provide in the OST program, how often do students:	SPP Staff				ACE-Only Staff			
	Never Occurs	Occurs Occasionally	Occurs Regularly	Always Occurs	Never Occurs	Occurs Occasionally	Occurs Regularly	Always Occurs
1. Work independently?	1.9%	35.2%	45.7%	17.1%	3.2%	52.4%	28.1%	16.2%
2. Use computer-based learning programs to complete activities and exercises?	15.5%	21.6%	32.5%	30.1%	25.5%	45.1%	15.2%	14.1%
3. Review and practice concepts taught during the school day?	6.1%	23.3%	40.9%	29.5%	15.6%	30.8%	35.1%	18.3%
4. Work individually with you on assigned tasks/activities?	2.8%	36.4%	41.7%	18.9%	7.6%	43.1%	34.9%	14.2%
5. Work with other students in small groups?	4.7%	20.3%	47.8%	27.0%	5.4%	27.3%	49.1%	18.0%
6. Choose what activities or projects they are going to work on or participate in?	13.9%	42.7%	35.0%	8.1%	9.2%	45.1%	30.9%	14.6%
7. Work on group projects that take more than one day to complete?	24.2%	37.6%	28.5%	9.5%	24.8%	41.0%	22.7%	11.3%
8. Listen to you deliver a lesson as part of whole group instruction?	10.8%	39.6%	34.4%	15.0%	12.9%	28.6%	36.2%	22.1%

In the activities <u>you</u> provide in the OST program, how often do students:	SPP Staff				ACE-Only Staff			
	Never Occurs	Occurs Occasionally	Occurs Regularly	Always Occurs	Never Occurs	Occurs Occasionally	Occurs Regularly	Always Occurs
9. Participate in interrelated activity sessions that build on one another to support specific skills?	6.1%	29.3%	46.9%	17.5%	11.4%	30.0%	38.2%	20.2%
10. Help or mentor other youth in completing a project or task?	14.7%	41.4%	32.8%	10.9%	15.6%	37.2%	31.8%	15.1%
11. Receive recognition for their work, achievements, or accomplishments?	2.8%	14.6%	48.1%	34.4%	6.6%	17.6%	46.4%	29.2%
12. Have the opportunity to set learning goals and monitor progress toward those goals?	9.9%	26.0%	45.4%	18.4%	16.3%	30.4%	34.7%	18.4%
13. Make choices about how they will engage with the content being covered in program offerings?	10.0%	33.3%	42.8%	13.8%	8.6%	30.4%	41.8%	19.0%
14. Have time to engage in reflection/summary at the end of the session to review what was learned?	7.1%	29.3%	44.0%	19.4%	11.9%	28.2%	39.6%	20.1%

Table C-5. *Internal Communication*, 2013–14

How frequently do you engage in the following tasks with other staff working in the OST program?	SPP Staff				ACE-Only Staff			
	Never	Once or Twice a Semester	About Once a Month	Nearly Every Week	Never	Once or Twice a Semester	About Once a Month	Nearly Every Week
1. Use program data for program planning	13.5%	17.5%	32.1%	36.6%	27.4%	16.9%	17.6%	38.0%
2. Have conversations about ways to improve programming	5.5%	10.5%	29.0%	55.0%	6.2%	18.8%	18.8%	55.9%
3. Work collaboratively to plan program activities or plan implementation of prescribed lessons	6.5%	10.5%	28.0%	55.0%	11.3%	11.9%	18.2%	58.4%
4. Collaborate on ways to increase student engagement	4.4%	8.8%	22.1%	64.5%	6.7%	14.0%	18.2%	60.9%
5. Work together to share ideas and approaches to effectively meet individual student needs	4.4%	10.2%	21.0%	64.2%	4.8%	11.5%	21.3%	62.1%
6. Discuss how program activities can better support students' academic needs	2.9%	12.8%	26.6%	57.6%	4.3%	17.5%	22.5%	55.6%
7. Participate in training and professional development	11.0%	38.0%	26.5%	24.5%	15.0%	38.5%	22.2%	24.1%
8. Identify research-based instructional practices to supplement existing programming and/or curriculum	8.1%	28.9%	27.4%	35.5%	24.1%	19.1%	29.0%	27.6%

Table C-6. *Program Climate*, 2013–14

Please rate the extent to which you agree or disagree with the following statements regarding the climate of the program:	SPP Staff				ACE-Only Staff			
	Strongly Disagree	Disagree	Agree	Strongly Agree	Strongly Disagree	Disagree	Agree	Strongly Agree
1. There is generally a spirit of cooperation within the OST program.	2.3%	0.9%	44.4%	52.1%	1.6%	2.2%	39.3%	56.7%
2. Employee morale is generally high at my OST program.	2.4%	6.7%	39.6%	51.2%	1.6%	2.8%	40.6%	54.8%
3. I enjoy working here.	2.8%	2.8%	28.7%	65.5%	2.1%	1.0%	23.0%	73.6%
4. I have adequate time and resources to plan activity sessions.	3.8%	12.0%	42.9%	41.0%	1.1%	6.2%	41.2%	51.4%
5. I have adequate time and resources to address individual student needs.	3.3%	17.7%	46.4%	32.5%	2.8%	8.5%	45.7%	42.8%
6. I am able to get technical assistance/support when I encounter challenges or problems.	3.4%	6.8%	43.8%	45.8%	2.3%	4.0%	40.4%	53.1%
7. Program leaders/management create a supportive, strength-based work environment.	2.8%	1.4%	39.9%	55.7%	1.0%	3.8%	30.6%	64.4%
8. I have timely access to program leaders/management to discuss issues that arise.	2.4%	2.4%	44.7%	50.4%	1.1%	2.7%	35.5%	60.5%
9. Program leaders/management are receptive to staff suggestions for designing and delivering program activities.	2.4%	1.4%	40.9%	55.1%	1.6%	1.6%	33.3%	63.2%

Table C-7. *Program Challenges*, 2013–14

Please indicate if you are experiencing any of the following types of challenges when leading OST activities, and if you are, how frequently these challenges occur.	SPP Staff				ACE-Only Staff			
	Never Occurs	Occurs Occasionally	Occurs Regularly	Always Occurs	Never Occurs	Occurs Occasionally	Occurs Regularly	Always Occurs
1. Low levels of student interest in program activities	20.4%	64.2%	12.3%	2.8%	28.6%	61.0%	9.7%	0.5%
2. Difficulty in effectively using curricular materials	54.5%	38.7%	4.3%	2.3%	58.9%	37.2%	3.7%	0.0%
3. Difficulty engaging students in curricular materials or program content	36.1%	53.8%	7.1%	2.8%	50.8%	43.2%	4.8%	1.0%
4. Difficulty in making content relevant to the lives of participating students	56.7%	37.0%	5.2%	0.9%	57.3%	36.0%	6.5%	0.0%
5. Lack of alignment between school and OST curriculum	64.7%	29.5%	4.2%	1.4%	69.2%	27.4%	2.7%	0.5%
6. Problems getting technology to work properly	49.7%	43.1%	5.2%	1.8%	62.6%	34.6%	2.1%	0.5%
7. Inconsistent program attendance	31.5%	51.1%	12.4%	4.7%	43.6%	48.6%	6.0%	1.6%
8. Instructional time interrupted by student behavioral issues	37.9%	47.8%	10.9%	3.3%	41.5%	50.2%	6.0%	2.1%

Appendix D. Analysis Methods

Rasch Analysis on Survey Responses

At its most basic level, the use of Rasch modeling techniques yields estimates of an individual respondent's ability and the relative difficulty of a given item appearing on the instrument in question (Bond & Fox, 2007). Working from the proposition that persons with greater ability will have a greater likelihood of successfully completing a given bank of test items (or find it easier to endorse survey items that demonstrate greater ability) than will less skilled persons, Rasch modeling techniques take person and item difficulty estimates yielded from an instrument, transform them by using a log function, and display them on a logit scale that allows person and item difficulties to be compared directly.

One of the benefits of using Rasch approaches is that they result in true interval-level scores that can be used when conducting analyses. In order to create true interval measures that could be employed effectively in supporting the domain of analyses needed for the report, we employed Rasch analysis techniques to create scale scores for scales associated with several instruments used to support data collection efforts during the 2012–13 and 2013–14 school years. Three different Rasch models were employed in this undertaking.

1. **Rasch Rating Scale Model (Linacre, 2005).** This model was used to calibrate scales appearing on the staff and student surveys and took the following form:

   $$\text{Log}(P_{nix} / P_{ni(x-1)}) = B_n - (D_i + R_x)$$

   where

   P_{nix} = the probability of person n of ability B_n being observed in category x of item i with difficulty D_i

   $P_{ni(x-1)}$ = the probability of person n of ability B_n being observed in category $x - 1$ of item i with difficulty D_i

   B_n = the ability of respondent n

   D_i = the difficulty of item i

   R_x = rating scale structure parameter for category x (indicates how much of the latent construction is covered by a given response category of the rating scale)

2. **Rasch Dichotomous Model (Wright & Masters, 1982).** This model was used to calibrate scales appearing on the APT-O and took the following form:

   $$\text{Log}(P_{ni} / (1 - P_{ni})) = B_n - D_i$$

   where

P_{ni} = the probability of activity *n* succeeding on item *i*

B_n = the ability of activity *n*

D_i = the difficulty of item *i*

3. **Many-Facet Rasch Measurement (Linacre & Wright, 2004).**– This model was used to calibrate scales appearing on the following observation instruments:

   a. PQA (both the Youth and School-Age versions of this instrument)
   b. OCES

   The many-facet Rasch measurement model employed in calibrating measures on the aforementioned instruments took the following form:

   $$\text{Log}(P_{nijk} / P_{nij(k-1)}) = B_n - D_i - C_j - F_k$$

   where

   P_{nijk} = the probability of activity *n* being given a rating of *k* on item *i* by rater *j*

   $P_{nij(k-1)}$ = the probability of activity *n* being given a rating of *k* - 1 on item *i* by rater *j*

   B_n = the ability of activity *n*

   D_i = the difficulty of item *i*

   C_j = the severity of rater *j*

   F_k = the difficulty of category *k* relative to category *k* -1

Quality and Student Engagement Correlational Analysis—2013–14

As noted earlier, during the spring semester of 2014, members of the evaluation team conducted two-day site visits at 26 sites—13 ACE centers with SPP funding and 13 ACE-only centers. During these visits, a total of 89 OST activities serving youth in Grades K–12 were observed by members of the evaluation team.

At the conclusion of each activity, both the PQA and the APT-O academic skill-building sections for reading, writing, and mathematics problem solving and communication were scored to obtain an estimate of activity-level quality. Nineteen of the 89 activities (approximately 21 percent) were observed by more than one member of the evaluation team. This method allowed the evaluation team to calibrate PQA and APT-O scores in a way in which systematic observer bias (some raters are inherently more severe in scoring the instruments, and others are inherently more lenient) could be quantified and adjusted for when determining a final activity score.

PQA Subscale Development

For this analysis, steps were taken to create subscale scores to explore better whether specific types of PQA practices were found to be related especially to youth engagement. In undertaking analyses to create subscale scores, some subscales demonstrated poor reliability, given a lack of variation in scores across activities (almost all activities scored highly on such subscales) and were therefore dropped from further analyses. Other subscales with poor reliability but good variability in scores were combined with other subscales to construct a usable scale. Ultimately, six subscale scores were calibrated from PQA data as shown in Table D-1.

Table D-1. PQA Subscales

New Subscale Name	Old Subscales That Make Up New Subscale
Activity Characteristics That Support Active Engagement	Activities support active engagement (*Supportive Environment* subscale H).
Encouragement and Feedback	Staff support youth in building new skills (*Supportive Environment* subscale I). Staff support youth with encouragement (*Supportive Environment* subscale J).
Cooperative Learning and Belonging	Youth have opportunities to develop a sense of belonging (*Interaction* subscale L). Youth have opportunities to collaborate (*Interaction* subscale M). Youth have opportunities to act as group facilitators and mentors (*Interaction* subscale N).
Positive Adult Interactions	Youth have opportunities to partner with adults (*Interaction* subscale O).
Planning and Choice	Youth have opportunities to set goals and make plans (*Engagement* subscale P). Youth have opportunities to make choices based on their interests (E*ngagement* subscale Q).
Reflection	Youth have opportunities to reflect (*Engagement* subscale R).

As in analyses conducted previously on observation scores for this evaluation, scores were placed on a scale from 0 to 100. The mean of these scales was then taken to create a total PQA score. Average scale scores for activities observed in SPP and ACE-only centers are outlined in Table D-2. Although some differences in the average PQA scores between activities in SPP and ACE-only centers are noticeable, none were significant.

Table D-2. Average PQA Scale Scores by Subscale and Center Type (2013–14)

PQA Subscales	Activities in SPP Centers (*n* = 45)	Activities in ACE-only Centers (*n* = 44)
Activity Characteristics That Support Active Engagement	65.069	64.030
Encouragement and Feedback	67.051	61.977
Cooperative Learning and Belonging	62.590	58.301
Positive Adult Interactions	59.504	56.707
Planning and Choice	40.678	48.634
Reflection	46.722	45.360
Total PQA Score	54.828	54.147

SOURCE: Texas 21st CCLC Evaluation, PQA Observation Scores, 2013–14.

APT-O scores were calibrated in a similar fashion, adjusted for observer bias and placed on a scale from 0 to 100. However, given that some activities were designed intentionally to support youth skill building in specific content areas like reading or mathematics and others had no content focus at all, a different approach was taken to classify activities on the basis of APT-O scores. Activities were classified first as having an intentional focus on building youth academic skills in reading, writing, or mathematics. Next, APT-O scale scores were created for the reading, writing, mathematical problem solving, and mathematical communication sections of the APT-O for each activity. Then, new variables were created for both reading and writing and for mathematics that indicated whether the activity both (1) was focused intentionally on building skills in that content area and (2) was above the 75th percentile on the APT-O scale score for that content area, indicating that multiple supports and opportunities for youth to build skills in that content area were provided by staff leading those activities. The number and percentage of activities observed in both SPP and ACE-only centers is outlined in Table D-3. Although some differences are present between SPP and ACE-only activities, none were significant.

Table D-3. Percentage of Activities by APT-O-Based Variables and Program Type (2013–14)

	Activities in Centers With SPP Programming (*n* = 45)		Activities in ACE-Only Centers (*n* = 44)	
APT-O–Based Variables	**Number**	**Percentage**	**Number**	**Percentage**
Intentionally Designed to Build Content Skills at or Above 75th Percentile on APT-O Content Scales—Reading and Writing	20	44.4%	13	29.5%
Intentionally Designed to Build Content Skills at or Above 75th Percentile on APT-O Content Scales—Mathematics	10	22.2%	11	25.0%

SOURCE: Texas 21st CCLC Evaluation, APT-O Observation Scores, 2013–14.

In order to explore the relationship between activity characteristics, quality, and levels of youth engagement, a series of analyses were conducted using engagement scores from both the OCES and the youth survey as outcomes. Each of the analyses used the variables related to activity characteristics outlined in Table D-4 and quality as predictors.

Table D-4. Characteristics of SPP and ACE-Only Activities (2013–14)

Characteristic	Definition
SPP Center	Indicated whether the center was funded by SPP or not
Academic Enrichment Activity	Indicated whether the activity being observed was classified as academic enrichment
Staff-to-Student Ratio	Represented the ratio of staff to youth participating in an observed activity
High Reading Content	Indicated whether the activity was (1) intentionally focused on building skills in reading and writing and (2) above the 75th percentile on the APT-O scale score for either reading or writing content
High Mathematics Content	Indicated whether the activity was (1) intentionally focused on building skills in mathematics and (2) above the 75th percentile on the APT-O scale score for either mathematics problem solving or communication
Total PQA Score	The mean scale score associated with the six subscales described previously in relation to the approach taken to score the PQA
PQA Subscales	Some analyses also included the six PQA subscales described previously.
Youth Grade Level	Analyses for which the OCES was the outcome of interest; included a variable that indicated if the observed activity was being attended by youth in Grades K–3. In cases in which the student survey engagement score was the outcome of interest, the grade level of the student survey respondent was included in the analysis.

HLM Analyses

On creating these characteristics, we ran a series of multiple regression analyses to explore how activity characteristics and quality were related to the OCES scores observers assigned to a given activity. Given that OCES scores were the only measure of engagement in relation to activities serving youth in Grades K–3, the multiple regression analyses were rerun but only with activities that served youth in Grades K–3.

In addition to looking at OCES scores, we conducted analyses related to the student surveys distributed at the end of each observed activity. A total of 401 student engagement surveys were collected during spring 2014 observations of 49 activities provided by both SPP and ACE-only programs represented in the sample. Given the nested nature of these data (youth nested within activities and activities nested within centers), a series of hierarchical linear models were run to explore the relationship

between activity-level characteristics and quality and the level of reported youth engagement in the activity session observed.

Initially, two hierarchical linear models were run. Both models treated the individual youth survey engagement score as the dependent variable of interest (i.e., the level one outcome), included youth grade level as a predictor at level one, and included the SPP status of the center as a level three predictor. The difference in the models related to the activity-level (i.e., level two) predictors that were included in the model. Model 1 examined the relationship between activity characteristics such as the activity being an academic enrichment offering or not and the staff-to-student ratio on the level of youth engagement, and Model 2 examined the relationship between PQA and APT-O–derived quality measures on youth engagement.

Model 1. Estimated Effect of Various Activity Characteristics on Student Survey Engagement Scores

Level 1 Model

$$Y = P_0 + P_1(GRADE\_LE) + E$$

Level 2 Model

$$P_0 = B_{00} + B_{01}(RATIO) + B_{02}(ACAD\_ENR) + R_0$$
$$P_1 = B_{10} + B_{11}(RATIO) + B_{12}(ACAD\_ENR)$$

Level 3 Model

$$B_{00} = G_{000} + G_{001}(TREAT) + U_{00}$$
$$B_{01} = G_{010} + G_{011}(TREAT) + U_{01}$$
$$B_{02} = G_{020} + G_{021}(TREAT) + U_{02}$$
$$B_{10} = G_{100} + G_{101}(TREAT)$$
$$B_{11} = G_{110} + G_{111}(TREAT) + U_{11}$$
$$B_{12} = G_{120} + G_{121}(TREAT) + U_{12}$$

Model 2. Estimated Effect of Various Quality Practices on Student Survey Engagement Scores

Level-1 Model

$$Y = P_0 + P_1*(GRADE\_LE) + E$$

Level-2 Model

$$P_0 = B_{00} + B_{01}(SW\_READ) + B_{02}(SW\_MATH) + B_{03}(PQA\_TOTA) + R_0$$
$$P_1 = B_{10} + B_{11}(SW\_READ) + B_{12}(SW\_MATH) + B_{13}(PQA\_TOTA)$$

Level-3 Model

$$B_{00} = G_{000} + G_{001}(TREAT) + U_{00}$$
$$B_{01} = G_{010} + G_{011}(TREAT) + U_{01}$$
$$B_{02} = G_{020} + G_{021}(TREAT) + U_{02}$$

$B_{03} = G_{030} + G_{031}(TREAT) + U_{03}$
$B_{10} = G_{100} + G_{101}(TREAT)$
$B_{11} = G_{110} + G_{111}(TREAT) + U_{11}$
$B_{12} = G_{120} + G_{121}(TREAT) + U_{12}$
$B_{13} = G_{130} + G_{131}(TREAT) + U_{13}$

One unexpected finding was that a significant, negative relationship was found between higher PQA scores and youth engagement in SPP programs. In order to explore this finding further, the evaluation team decided to rerun the model but without the inclusion of the variable indicating the SPP status of the center.

Model 3. Estimated Effect of Various Quality Practices on Student Survey Engagement Scores Not Controlling for SPP Status

Level-1 Model

$Y = P_0 + P1*(GRADE\_LE) + E$

Level-2 Model

$P0 = B_{00} + B_{01}(SW\_READ) + B_{02}*(SW\_MATH) + B_{03}(PQA\_TOTA) + R_0$
$P1 = B_{10} + B_{11(}SW\_READ) + B_{12}*(SW\_MATH) + B_{13}(PQA\_TOTA)$

Level-3 Model

$B_{00} = G_{000} + U_{00}$
$B_{01} = G_{010} + U_{01}$
$B_{02} = G_{020} + U_{02}$
$B_{03} = G_{030} + U_{03}$
$B_{10} = G_{100}$
$B_{11} = G_{110} + U_{11}$
$B_{12} = G_{120} + U_{12}$
$B_{13} = G_{130} + U_{13}$

Finally, in order to explore the role individual practices specified in the PQA may have in supporting youth engagement, a series of models were run in which the only activity-level predictor included in the model was one of the six PQA subscales described previously. These models also included student grade level (at level one) and SPP status of the center (at level three).

Level-1 Model

$Y = P0 + P1*(GRADE\_LE) + E$

Level-2 Model

$P0 = B00 + B01*(QUALITY\ PRACTICE) + R0$
$P1 = B10 + B11*(QUALITY\ PRACTICE)$

Level-3 Model

B00 = G000 + G001(TREAT) + U00
B01 = G010 + G011(TREAT) + U01
B10 = G100 + G101(TREAT)
B11 = G110 + G111(TREAT) + U11

Student Survey Pre-Post Differences

2012–13

To examine the impact of SPP participation, we examined the academic Mindsets and Behaviors Survey and activity leader survey scale scores to assess differences between preadministration and postadministration. Specifically, the analysis compared the performance of students who participated in SPP with that of similar students who participated only in ACE-only program activities.

In any evaluation of a program in which participants are not assigned randomly to participate in the program, the problem of selection is paramount. It is likely that students who participated in SPP programming were different from those who did not participate. These 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, SPP students began the program year with lower scores on the survey constructs than did their peers who participated in ACE-only programming. This reflects the fact that the selected population for the SPP programs was more at risk academically than were the students selected for ACE programs.

To mitigate any existing bias related to program effect that might emerge by simply comparing SPP students with ACE-only students, propensity score matching (PSM) was employed. PSM is a quasi-experimental approach whose two-stage process is designed to address the problem of existing bias. In the first stage, the probability that each student will participate in the SPP program was modeled on observable characteristics captured in data received from TEA. These characteristics were used to compare participating and nonparticipating students who would have had a similar propensity to select into the program. In the second stage, the predicted probability of participation was used to model student outcomes while accounting for selection bias. Each stage is described in greater detail in the following sections.

Stage 1: Creation of the Comparison Group. The comparison group was identified by first identifying a student's treatment status as either an SPP or ACE-only participant. Then several student-level variables were used to identify ACE-only participants who had characteristics similar to those of SPP participants, including the following:

- Prior achievement in reading and mathematics
- Prior scale scores on the four constructs
- Student demographic information including the following:

- Gender
- Racial status
- Socioeconomic status
- Special education status
- Migrant status
- Immigrant status
- English proficiency level
- Gifted status

Demographic differences are presented in Tables D-5 and D-6.

Table D-5. Grade K–3 Characteristics (2012–13)

	SPP	ACE-Only	Matched Comparison
Student-Level Characteristics			
Number of absences 2011–12	5.26	4.13	4.12
Number of disciplinary incidents 2011–12[1]	1.55	2.14	1.60
Female	50%	58%	57%
Male	50%	42%	43%
Special education	7%	4%	4%
Limited English proficiency	53%	48%	47%
[1] This includes only students with at least one disciplinary incident. *N* = 769			

SOURCE: TEA, PEIMS, 2011–12 and 2012–13.

Table D-6. Grade 4–12 Characteristics (2012–13)

	SPP	ACE-Only	Matched Comparison
Student-Level Characteristics			
Number of absences 2011–12	4.87	4.18	4.22
Number of disciplinary incidents 2011–12[1]	2.65	2.06	1.84
Female	50%	47%	44%
Male	50%	53%	56%
Special education	11%	9%	10%
Limited English proficiency	40%	25%	36%
[1] This includes only students with at least one disciplinary incident. *N* = 1,102			

SOURCE: TEA, PEIMS, 2011–12 and 2012–13.

In addition to the student-level variables, the propensity score model included school variables that added information about the school a student attended. This accounted

for school-based contextual factors, which might have accounted for differences in the propensity for a student to participate. A total of 111 variables were considered for the propensity score model. The propensity score model was fit separately for the two grade levels, K–3 and 4–12.

The final propensity score models for the two grade levels were checked to ensure that the analysis sample was balanced across relevant covariates. The models produced comparison samples (ACE-only students), which were balanced with the treatment samples (SPP students) across variables that were important for predicting a student's enrollment in SPP. As measured by these variables, the result indicates that the treatment and comparison groups resulting from the matching model were not significantly different from one another prior to treatment.

*Stage 2: Statistical Modeling of Student Outcomes***.** Outcomes of students in the SPP program then were compared with the outcomes of ACE-enrolled students. We balanced pretreatment group differences in observed covariates by using a propensity score stratification and marginal mean weighting approach (Hong & Hong, 2009). Various strata were used based on the spread and overlap of the data. The propensity score logit along with the pretreatment measure of the outcome also were included in the outcome model to control for within-strata differences and residual bias (Schafer & Kang, 2008). Student outcomes were modeled using two-level hierarchical linear models to account for the nested nature of the data (students within schools) as follows:
Level 1 – Students

$$y_{ij} = \beta_{0j} + \beta_{1j}\, Participation_{ij} + \sum_{S=2}^{15} \beta_s L_{sij} + \beta_{16j} LP_{ij} + \beta_{17j} Pretest_{ij} + r_{ij}$$

Where y_{ij} is the student level outcomes (survey subscales), $Participation_{ij}$ is an indicator of whether the student participated in the SPP program, L_{sij} is an indicator variable for the logit propensity score stratum, LP_{ij} is the logit propensity score, and $Pretest_{ij}$ is the pretreatment measure of the outcome. Subscripts *i*, *j*, and *s* correspond to student, school, and strata, respectively.

Level 2 – Campus

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

The Level 2 equation includes only β_{0j} because the hierarchical linear model is a random intercept model; all other coefficients (i.e., participation indicator, logit propensity score stratum, logit propensity score, and pretreatment indicator) at Level 1 are fixed and, therefore, not listed at Level 2. Because the treatment and comparison groups were matched using all of the covariates described earlier, it is not necessary to include these variables in the final outcome model.

Steps also were taken to compare survey outcomes for Grade K–3 centers adopting either a *Learning Strategies* or *Skills-Building* approach. HLM was used to assess the relationship between the program model adopted by the SPP center and youth

outcomes collected from the survey. The following HLM model was run in undertaking this analysis.

The two-level model of correlation between program participation and student performance (written in mixed model format) is as follows:

$$Y_{ij} = \beta_0 + \beta_1 StudentCovariate_{ij} + \beta_2 Treatment_{ij} + u_j + r_{ij}$$

where

- Y_{ij} is the performance of student *i* in school *j*;
- β_0 is a constant term showing average student performance in the comparison group;
- $StudentCovariate_{ij}$ is a vector of student-level covariates, including their prior performance on the outcome of interest, gender, ethnicity, economically disadvantaged status, limited English proficiency status, number of years enrolled in the program, and total days of SPP program participation;
- β_1 is a vector of coefficients associated with each of those covariates showing the association of each student-level characteristic and the outcome;
- $Treatment_{ij}$ is the treatment status for student *i* in school *j*;
- β_2 shows the average difference in performance between treatment group and comparison group;
- u_j is a school-level random error term, with an assumed normal distribution with mean zero and variance τ; and
- r_{ij} is an student-level error term, also assumed to have a normal distribution with mean zero and variance σ^2.

2013–14

Although the methodology was different during 2013–14 in that the evaluation team did not employ a PSM process, it is still useful to understand the demographic characteristics of the participants taking the survey in 2013–14. This allows us to understand how similar or different they were from the previous year's participants and how different the two types of students are from one another. Tables D-7 and D-8 show the demographic characteristics of students included in the 2013–14 analysis.

Table D-7. Grade K–3 Characteristics (2013–14)

	SPP	ACE-Only
Student-Level Characteristics		
Female	46%	54%
Male	53%	46%
Special education	9%	7%
Limited English proficiency	40%	34%
N = 1,478		

SOURCE: TEA, PEIMS, 2013–14.

Table D-8. Grade 4–12 Characteristics (2013–14)

	SPP	ACE-Only
Student-Level Characteristics		
Female	47%	53%
Male	53%	7%
Special education	8%	9%
Limited English proficiency	28%	37%
N = 2,426		

SOURCE: TEA, PEIMS, 2013–14.

Because most of students took both the pretest (fall 2013) and posttest (spring 2014) surveys, a paired sample *t*-test was used to check the difference between the two administrations. Steps were taken to examine and report on the descriptive statistics of the construct scale scores of the two administrations as well as look at the significance of the mean difference between the two administrations. Because it is anticipated that more rigorous statistical techniques will be used in the 2014–15 evaluation to examine the relationship between programming and change on the Mindsets and Behaviors Survey, the 2013–14 analysis intentionally was limited to looking strictly at the mean differences and reporting on descriptive information.

Appendix E. Survey Constructs

Activity Leader Survey Constructs

Table E-1. Items Included in the Organizational Quality Constructs (2012–13 and 2013–14)

Internal Communication	***Program Climate***
How frequently do you engage in the following tasks with other staff working in the Out-of-School time (OST) program... Use program data for program planning Have conversations about ways to improve programming Work collaboratively to plan program activities or plan implementation of prescribed lessons Collaborate on ways to increase student engagement Work together to share ideas and approaches to effectively meet individual student needs Discuss how program activities can better support students' academic needs Participate in training and professional development Identify research-based instructional practices to supplement existing programming and/or curriculum	*Please rate the extent to which you agree or disagree with the following statements regarding the climate of the program...* There is generally a spirit of cooperation within the OST program. Employee morale is generally high at my OST program. I enjoy working here. I have adequate time and resources to plan activity sessions. I have adequate time and resources to address individual student needs. I am able to get technical assistance/support when I encounter challenges or problems. Program leaders/management create a supportive, strength-based work environment. I have timely access to program leaders/management to discuss issues that arise. Program leaders/management are receptive to staff suggestions for designing and delivering program activities.
Response Options: never, once or twice a semester, about once a month, and nearly every week	Response Options: strongly disagree, disagree, agree, and strongly agree

Table E-2. Items Included in the Planning Constructs (2012–13 and 2013–14)

Access to Student Data	***Intentional and Responsive Instruction***
Please indicate whether you receive each of the following, and to what extent you use it in planning for the activities you provide... Individual student academic plans Students' state assessment scores Students' scores on district- or building-level assessments Students' grades Input from students' school-day teachers Other	*How often do you lead or participate in out-of-school-time (OST) program activities that are...* Based on written lessons outlining the specific purposes of a given activity session, assignment, or project Designed to support specific, measurable learning goals for students Planned in advance Meant to extend and enhance skills developed in prior activity sessions or lessons Explicitly meant to promote skill building and mastery in relation to one or more state or district standard(s) Discussed and planned with other ACE/SPP staff prior to activity session delivery Discussed and planned with other school-day staff prior to activity session delivery Informed by levels of student engagement in previous activity sessions covering similar content Intentionally designed to address prior student mistakes or misunderstandings
Response Options: do not receive, occasionally use, and often use	Response Options: never occurs, occurs occasionally (some sessions), occurs regularly (most sessions), and always occurs (all sessions)

Table E-3. Items Included in the Instructional Constructs (2012–13 and 2013–14)

Individualized and Contingent Instruction	***Academic Learning Activities***
Please indicate the degree to which you know the following about the students you work with when providing OST programming… I usually know if a student is struggling with academic content during the regular school day. I know how a student performed on district- or building-level assessments. I am generally aware of the specific academic concepts and skills a student is struggling with. I understand the learning styles of each student I work with. I am aware of the strengths of each student. I recognize areas a student needs to improve in. I have a good understanding of how a student feels about school. I know how a student's prior experiences with academic content/concepts should impact individual instruction with that student.	*In the activities you provide in the OST program, how often do students...* Work independently Use computer-based learning programs to complete activities and exercises Review and practice concepts taught during the school day Work individually with you on assigned tasks/activities Listen to you deliver a lesson as part of whole group instruction
Response Options: I know this for none of my students, I know this for some of my students, I know this for most of my students, and I know this for all of my students.	Response Options: never occurs, occurs occasionally (some sessions), occurs regularly (most sessions), and always occurs (all sessions)

Table E-4. Items Included in the *Youth Development Activities* Construct (2012–13 and 2013–14)

In the activities you provide in the afterschool program, how often do students... Work with other students in small groups Choose what activities or projects they are going to work on or participate in Work on group projects that take more than one day to complete Participate in interrelated activity sessions that build on one another to support specific skills Help or mentor other youth in completing a project or task Receive recognition for their work, achievements, or accomplishments Have the opportunity to set learning goals and monitor progress toward those goals Make choices about how they will engage with the content being covered in program offerings Have time to engage in reflection/summary at the end of the session to review what was learned
Response Options: never occurs, occurs occasionally (some sessions), occurs regularly (most sessions), and always occurs (all sessions)

Table E-5. Items on the *Program Challenges* Construct (2012–13 and 2013–14)

Please indicate if you are experiencing any of the following types of challenges when leading OST activities, and if you are, how frequently these challenges occur. Low levels of student interest in program activities Difficulty in effectively using curricular materials Difficulty engaging students in curricular materials or program content Difficulty in making content relevant to the lives of participating students Lack of alignment between school and OST curriculum Problems getting technology to work properly Inconsistent program attendance Instructional time interrupted by student behavioral issues Other
Response Options: never occurs, occurs occasionally (some sessions), occurs regularly (most sessions), and always occurs (all sessions)

Student Engagement Survey Construct

Table E-6. Student Engagement Survey Items (2012–13 and 2013–14)

Student Engagement Survey Items
Thinking about today's activity that just ended, please answer the following questions: Was today's activity interesting? Was this activity important to you? Did you really have to focus to do the activity? Did you enjoy what you were doing during this activity? Was it easy to pay attention during today's activity? Was the activity something you were good at doing? Did you wish you were doing something else? Did you feel like you had a say in what you did during the activity?
Response Options: not at all, sort of, and very much

Academic Mindsets and Behaviors Survey Constructs

Grades K–3

Table E-7. Activity Leader Survey Items Related to *Effort and Persistence* and to *Learner Behaviors* (2012–13)

***Effort and Persistence* (7 Items)**	***Learner Behaviors* (10 Items)**
The student: Focuses on and completes schoolwork Is confident about his/her ability to do well in school Tries to complete schoolwork even if it is difficult Keeps working on assignments/activities when frustrated Works really hard in the program Feels like they are "smart enough" to do well in school Works hard to understand and overcome academic concepts they struggle with	*The student:* Is alert and focused when learning Is interested in learning new things Uses time wisely to accomplish work Seeks out information and resources needed to accomplish work Actively participates in activities and lessons Is able to work independently when necessary Monitors own learning and understanding Focuses on understanding work and assignments Seeks help from others when confused about work or assignments Recognizes areas/concepts he/she struggles with and seeks appropriate help in a timely manner
Response Options: not at all true, not very true, sort of true, and very true	

Table E-8. Activity Leader and Teacher Survey Items Related to *Effort and Persistence* and to *Learner Behaviors* (2013–14)

***Effort and Persistence* (6 Items)**	***Learner Behaviors* (7 Items)**
The student: Focuses on and completes schoolwork Is confident about his/her ability to do well in school Tries to complete schoolwork even if it is difficult Keeps working on assignments/activities when frustrated Feels like they are "smart enough" to do well in school Works hard to understand and overcome academic concepts they struggle with	*The student:* Is alert and focused when learning Uses time wisely to accomplish work Seeks out information and resources needed to accomplish work Is able to work independently when necessary Monitors own learning and understanding Focuses on understanding work and assignments Recognizes areas/concepts he/she struggles with and seeks appropriate help in a timely manner
Response Options: not at all true, not very true, sort of true, and very true	

Academic Performance **Items (Completed by School-Day Teacher) (5 Items)**
1. Reading a. Beginning reading/reading skills b. Reading comprehension c. Writing 2. Mathematics a. Number, operations, and quantitative reasoning b. Underlying processes and mathematical tools (e.g., problem solving)
Response Options: poor (significantly below grade-level standards), needs improvement (somewhat below grade-level standards), satisfactory (at grade-level standards), and very good (above grade-level standards)

Grades 4–12

Table E-9. Survey Items Related to *Effort and Persistence* and *Learner Behaviors* (2012–13)

Effort and Persistence **(6 Items)**	***Learner Behaviors*** **(7 Items)**
I try to do my schoolwork even when it looks really hard. When I am taught something that doesn't make sense, I keep trying to figure it out. I keep trying to do my schoolwork even if it is hard. I work really hard in school. I don't give up on my schoolwork even when I am frustrated. I try harder when I don't understand.	Before I start my schoolwork, I make sure I have all the things I need. I use my time in class to do my work and keep up with the rest of the class. I usually take part in what we do in class. When I'm in class, I think about what we are working on. I listen carefully in class. I am interested in the things we work on in class. I think most of my classes are fun.
Response Options: not at all like me, sort of like me, and a lot like me	

Table E-10. Survey Items Related to *Self-Competence* and *Mastery Orientation* (2012–13)

Self-Efficacy/Self-Competence **(6 Items)**	***Mastery Orientation*** **(7 Items)**
I can do my schoolwork well if I try hard. I'm pretty smart in school. I can get good grades in school. I just have to work hard to do well in school. I can do schoolwork that is hard as long as I don't give up. If I don't do well on my schoolwork, it's because I didn't try hard enough.	I always try to figure things out in class. I look at my schoolwork when it is returned to see what I got right and wrong. I try to understand my schoolwork. I like to know the answer to a question, even if it is not on a test. I think about what I am going to learn before I start my schoolwork. I want to understand my schoolwork well. I try to get mostly A's and B's in school.
Response Options: not at all like me, sort of like me, and a lot like me	

Table E-11. Survey Items Related to *Effort and Persistence* and *Learner Behaviors* (2013–14)

***Effort and Persistence* (6 Items)**	***Learner Behaviors* (7 Items)**
I try to do my schoolwork even when it looks really hard. When I am taught something that doesn't make sense, I keep trying to figure it out. I keep trying to do my schoolwork even if it is hard. I work really hard in school. I don't give up on my schoolwork even when I am frustrated. I try harder when I don't understand.	Before I start my schoolwork, I make sure I have all the things I need. I use my time in class to do my work and keep up with the rest of the class. I usually take part in what we do in class. When I'm in class, I think about what we are working on. I listen carefully in class. I am interested in the things we work on in class. I think most of my classes are fun.
Response Options: not at all like me, sort of like me, and a lot like me	

Table E-12. Survey Items Related to *Sense of Competence as a Learner* (2013–14)

***Sense of Competence as a Learner* (5 Items)**	***Sense of Competence as a Reader* (5 Items)**
I like to give new things a try, even if they look hard. In school, I'm as good as other kids. I'm good at solving problems. I'm as good as other kids my age at learning new things. When I can't learn something right away, I keep trying until I get it.	I like to read at home during my free time. I enjoy reading when I'm at school. I enjoy reading when I'm at this afterschool program. I'm good at reading. I like to give new books a try, even if they look hard.
***Sense of Competence in Mathematics* (6 Items)**	
I like to learn new things in math. I like to do math when I'm at school. I like to do math when I'm at this afterschool program. Math is something I'm good at. I'm interested in math. I like to give new math problems a try, even when they look hard.	
Response Options: not at all like me, sort of like me, and a lot like me	

Appendix F. Examples of SPP Activities by Type

To following are examples that illustrate the different types of SPP activities. The primary examples have relatively high PQA student engagement scores (among the top 50 percent).

Face-to-Face Format, *Learning Strategies* Approach

The SPP program at Center 1 serves elementary students and was classified as a *Learning Strategies* program. The observed activity received the second highest score on engagement and the highest rating on the interaction scale. The SPP program uses the Principles of Learning curriculum, which was developed by the UTeach program at the University of Texas. The curriculum emphasizes inquiry and critical thinking and is designed to encourage engagement in learning. The SPP program activity leaders at Center 1 were UTeach students who were supported by certified mentor teachers.

The observed activity was one in a series of lessons that would result in students cooking a traditional Mexican meal for a fiesta. In the observed activity, elementary students first participated in a logic puzzle to identify different combinations of food dishes that could be made from six different Mexican dishes. This was followed by a mathematics exercise in which students calculated the cost of each of the food combinations and then decided which to include in a two-dish meal. Students used handheld whiteboards and markers to show their work to teachers. After choosing their meal, students had to determine whether they had the money (coins in baggies that students counted) to pay for the meal they had selected. Students wrote in journals as they answered questions posed by the activity leader and were required to use full sentences during the journaling exercise. After this activity, students played a Mad Lib game during which one student selected the words and the partner read the story out loud. Students then switched roles.

The UTeach student teacher paid close attention to each student's progress. If some students completed a portion of the exercise before others, he gave them additional, more complex work within the same module. When all students completed a particular module of the lesson, they moved onto the next activity as a group.

According to the interview respondents, each Friday, UTeach student teachers and the University of Texas lesson developers discuss the classes and identify whether the lesson plans should be modified. Student teachers are observed by UTeach staff, site coordinators, and the mentor teachers and receive regular constructive feedback.

The example from Center 1 is one of a highly structured program model. Another SPP activity that had a *Learning Strategies* approach was a project-based learning activity that did not follow a specific curriculum. This activity was rated the lowest on the PQA student engagement scale. The observed activity began with a short video on photovoltaics, the subject of a project-based learning sequence of lessons. When the video ended, students worked in groups to calculate the amount of energy saved by use of photovoltaics and then moved to small groups to build solar cars. The session was

staffed with five teachers who moved from group to group to provide assistance. One student group struggled with the worksheet and was unable to stay on task (leading to some observable disruption during the session). Staff frequently intervened with this group. During the solar car-building phase of the session, some groups (including a student who chose to work independently) finished very quickly but did not have another activity in which to engage.

Computer-Based and Face-to-Face Format, *Learning Strategies* Approach

Center 2 uses a CGI curriculum, which equips students with a variety of mathematics problem-solving strategies. Metacognition is an important part of the program, and students were asked to think about their thinking by answering questions about why they determined their answer was right or wrong. Students were in the early elementary grades. Among the 22 observed activities, the CGI activity had the highest rating on student engagement.

During the CGI observed activity, students were encouraged to try new approaches for solving problems and were asked to articulate the strategy they used. Half of the class worked on the computer-based Essential Skills mathematics program with headsets on, and the other half of the class worked with the teacher on mathematics problems by using CGI strategies. The computer-based component, Essential Skills, used a video game that incorporated skills and concepts associated with STAAR Category 1 (numbers, operations, quantitative reasoning) and Category 2 (patterns, relationships, and algebraic reasoning). Students also used a computer-based Mentoring Minds mathematics program for assessment purposes.

The face-to-face component of the CGI session began with all students going around the room sharing something that they learned that day. The activity provided opportunities for students to take a leadership role and explain their problem-solving strategies to the small group of students. The lead activity leader for the CGI session had been trained on the CGI approach and had several years of experience using CGI in the regular school-day setting. She was a highly skilled activity leader and was chosen for the SPP program because of her experience, skills, and excellent rapport with students. Student engagement was high for both the computer-based and face-to-face components but was higher for the face-to-face CGI activity.

Computer-Based and Face-to-Face Delivery Mode, *Skills-Building* Approach

The SPP intervention at Center 3 uses the Marie Carbo framework in its Power Reading session. Power Reading is a computer-based program in which lessons are based on a pretest that establishes each student's reading level. During the Power Reading session that was observed, students went through several steps. First, through their headsets, they listened to a passage being read. They then listened again, this time reading along with the passage. Students then played a game, which was designed to test their comprehension of the passage. Following the game, students read the passage again.

When students finished the individual work on the computer, they indicated to the activity leader that they were ready to do a "read back," during which they read the passage in front of the teacher who listened for errors, fluency, and inflections. Students then were asked to demonstrate their comprehension by summarizing the passage to the teacher. During the teacher-student interactions, students were provided with a progress report, and the teacher explained how much they had progressed as readers during the course of the program. The feedback session between the teacher and student was a powerful aspect of the session, which involved positive reinforcement of the progress each student was making in the program. This kind of reinforcement can help to improve a student's academic mindset and sense of competence, which can, as posited in the SPP theory of change, lead to generalized school success. The students were provided with a printout of their progress to bring home to their parents.

The Power Reading activity leaders were selected, according to the project directors and site coordinator, through "a very competitive" districtwide process during which "only the very best interventionists were selected for the program."

Computer-Based Format, *Skills-Building* Approach

A computer-based-only session was observed at Center 4. The SPP program uses Study Island, a computer-based reading intervention program. During the observed activity, students worked on reading passages selected by the activity leader who based reading selections on benchmark scores from previous sessions. Students sat at computers with headphones on, read passages, and then answered questions related to the passages. Students had to answer seven questions correctly before being able to play a game the system sets up for them (an incentive for most of the students). The skills students addressed in the Study Island session were ones that were taught during the school day.

The teacher in the observed activity consistently interacted with the students, providing direct support and assistance as needed. Students expressed excitement when they completed a Study Island module successfully and attained a score higher than their last. Conversely, students who had not improved their scores expressed frustration. However, the teacher was successful at redirecting the students through extremely positive encouragement.

Center 5 uses a computer-based format in which students work on the Achieve3000 computer learning program. This is another program in which students' reading levels are established and then students are assigned articles at the appropriate levels. In the observed sessions, two certified teachers were responsible for six students each. At the beginning of the session, the teachers sent students an e-mail with a guiding question for the day. The teachers frequently interacted with students through an electronic messaging system that was integrated into the software program, with students asking questions and teachers prompting students to think about a topic. Teachers also frequently circulated throughout the room and interacted with students in a way that was consistently warm and encouraging.

Face-to-Face Format, *Skills-Building* Approach

One of the observed activities using a face-to-face format and *Skills-Building* approach was an academic enrichment session observed at Center 6. The observed session was a component of the SPP program's weekly "Flight School" and was led by an external vendor, Phoenix Arising. Students learned about aviation from the perspective of pilots and aviation engineers. Students assigned to the pilot group participated in a computer-based flight simulation exercise, and those assigned to the engineer group engaged in hands-on activities that demonstrated aviation terms and physics concepts, such as force. The second hour of the enrichment session focused on journaling—students writing down what they learned during the previous session. The journaling session was led by SPP staff at Center 6 who had received training at Phoenix Arising. Students were far more engaged in the activities led by the activity leader than in the journaling exercise.

Another academically infused enrichment session, offered at Center 2, incorporated mathematics skills into a college and career readiness activity. This session did not have benchmark assessments but did incorporate activity leader and student question-and-answer methods that provided students with immediate feedback on their understanding. This session also provided students with a large degree of choice in the activities.

Appendix G. PQA Scores for Observed SPP Activities, 2012–13

Observation	Approach	Mode	Grade Level	Environment	Interaction	Engagement	Total
1	LS	CB, FF	ES	90.69	72.00	57.53	73.41
2	LS	FF	ES	80.37	80.63	55.16	72.05
3	LS	FF	ES	74.20	65.74	55.16	65.03
4	SB	CB, FF	MS	71.91	57.87	52.20	60.66
5	SB	FF	ES	74.20	60.49	48.65	61.11
6	SB	CB, FF	ES	74.11	55.75	48.65	59.50
7	SB	CB, FF	ES	64.21	46.79	48.65	53.22
8	SB	FF	ES	48.22	17.01	31.56	54.57
9	SB	FF	ES	61.51	55.75	46.45	43.77
10	LS	FF	MS	67.17	18.61	45.52	65.34
11	SB	FF	ES	80.37	71.32	44.33	50.08
12	SB	FF	ES	58.63	50.93	40.69	45.80
13	SB	CB	MS	60.74	36.38	40.27	45.52
14	SB	FF	ES	59.90	41.96	34.69	39.17
15	SB	CB	ES	45.52	37.31	34.69	32.26
16	SB	CB	ES	51.61	37.65	30.88	43.91
17	SB	FF	ES	64.89	35.96	30.88	40.05
18	SB	CB	HS	60.74	30.88	30.20	40.61
19	SB	CB	MS	60.07	35.96	27.75	41.26
20	SB	FF	MS	69.54	46.36	17.34	44.41
21	SB	CB	HS	67.34	35.96	17.34	40.21
22	SB	CB	ES	52.54	29.36	17.09	33.00
23	LS	FF	HS	61.76	39.59	16.92	39.42
			Average	65.23	46.10	37.93	49.75

Observation	Approach	Mode	Grade Level	Environment	Interaction	Engagement	Total
LS = Learning Strategies, SB = Skills-Based, CB = Computer-based, FF = Face-to-Face, ES = Elementary School, MS = Middle School, HS = High School.							

Appendix H. Exploring the Validity and Reliability of the Academic Mindsets and Behaviors Surveys

Both the student survey for students in Grades 4–12 and the activity leader reports for students in Grades K–3 were analyzed psychometrically by using a computer program that tests both the reliability and the internal consistency of survey instruments. Reliabilities range from 0 to 1, with values closest to 1 generally being considered best, as shown in Table H-1.

Table H-1. Internal Consistency Associated With Reliability Statistics

Reliability Statistic (x)	Internal consistency
$x \geq 0.9$	Excellent
$0.8 \leq x < 0.9$	Good
$0.7 \leq x < 0.8$	Acceptable
$0.6 \leq x < 0.7$	Questionable
$0.5 \leq x < 0.6$	Poor
$x < 0.5$	Unacceptable

Although the guidelines presented in Table H-1 are general, the criterion for an acceptable reliability depends on the intended use of the scores. If the scores carry high stakes (i.e., for individual students), then higher reliabilities might be required. However, if the intent is to use these scores as an outcome and model group differences, low reliabilities are not overly problematic— other than lowering the power of the statistical test.

Activity Leader Reports (Grades K–3)

In 2012–13, the two scales in the activity leader reports functioned well in the two administrations, although the psychometric analysis identified some areas for improvement, as shown in Table H-2. For both scales, *Effort and Persistence* and *Learner Behaviors*, the reliability scores are in the *good* threshold or higher (0.8 or above). However, in both administrations, the analysis results show that there is a "ceiling" effect with the two constructs—that is, many activity leader ratings of individual students bunched at the top of the scales (leaders reported many high abilities and few items with high difficulties). In addition, the items are generally easy to agree with.

Table H-2. Reliability Statistics by Construct on Activity Leader Reports, Grades K–3 (2012–13)

	Preadministration		Postadministration	
Construct	Cronbach's α	Rasch Reliability	Cronbach's α	Rasch Reliability
Effort and Persistence	0.97	0.84	0.96	0.85
Learner Behaviors	0.97	0.85	0.98	0.86
N = 821				

SOURCE: 21st CCLC Evaluation, Grade K–3 Activity Leader Mindsets and Behaviors Report, 2012–13.

On the basis of these analyses, the evaluation team revised the activity leader survey items slightly. Several items were revised to make them more difficult to endorse or removed all together to attempt to improve the ceiling effects of the constructs overall. In addition, a new construct related to *Academic Performance* was included to begin to get at the impact of the programs on academics. The reliability scores are again in the *good* threshold or higher, as shown in Table H-3.

Table H-3. Reliability Statistics by Construct on Activity Leader Reports, Grades K–3 (2013–14)

	Preadministration		Postadministration	
Construct	Cronbach's α	Rasch Reliability	Cronbach's α	Rasch Reliability
Effort and Persistence	0.96	0.86	0.96	0.84
Learner Behaviors	0.96	0.88	0.96	0.87
Academic Performance	0.95	0.91	0.95	0.90
N = 779				

SOURCE: 21st CCLC Evaluation, Grades K–3 Activity Leader Mindsets and Behaviors Report, 2013–14.

Academic Mindsets and Behaviors Survey (Grades 4–12)

When compared with the activity leader survey, the academic Mindsets and Behaviors Survey had more room for improvement, though overall it functioned as intended after its first administration in Year 3 of the evaluation. As shown in Table H-4, during Year 3, reliabilities for the survey overall ranged from 0.53 to 0.74, which are in the *poor* to *acceptable* range. During postadministration, reliabilities on two of the four constructs were in the *acceptable* threshold or higher. However, in preadministration, reliabilities of all four constructs were below the *acceptable* range.

Table H-4. Reliability Statistics by Construct on Skills and Beliefs Survey, Grades 4–12 (2012–13)

	Preadministration		Postadministration	
Construct	Cronbach's α	Rasch Reliability	Cronbach's α	Rasch Reliability
Effort and Persistence	0.83	0.68	0.86	0.74
Learner Behaviors	0.81	0.69	0.84	0.73
Self-Efficacy/Self-Competence	0.64	0.53	0.72	0.59
Mastery Orientation	0.77	0.61	0.81	0.68
N = 1,175				

SOURCE: 21st CCLC Evaluation, Grades 4–12 Mindsets and Behaviors Survey, 2012–13.

On the basis of this analysis, the survey was revised considerably. The *Effort and Persistence* and *Learner Behavior* constructs remained the same because those scales performed better than the other two scales in 2012–13. As Table H-5 illustrates, the reliability ratings remained roughly the same. In both cases, the constructs performed in the just barely acceptable range (0.7 or above). The *Self-Efficacy/Self-Competence* and *Mastery Orientation* constructs were removed from the survey altogether because of their poor performance ranging from *poor* to *questionable* in 2012–13. These were replaced with new scales related to *Sense of Competence*. These scales were borrowed from existing validated tools and perform at much higher levels. As Table H-5 shows, the *Sense of Competence as a Learner* scale is the only one performing below the acceptable level, but only by a small margin. The other two constructs performed in the acceptable to good range. This finding suggests these constructs were an improvement on the two constructs that were removed.

Table H-5. Reliability Statistics by Construct on Skills and Beliefs Survey, Grades 4–12 (2013–14)

	Preadministration		Postadministration	
Construct	Cronbach's α	Rasch Reliability	Cronbach's α	Rasch Reliability
Effort and Persistence	0.85	0.71	0.88	0.74
Learner Behaviors	0.83	0.69	0.84	0.72
Sense of Competence as a Learner	0.77	0.64	0.79	0.66
Sense of Competence as a Reader	0.81	0.73	0.85	0.75
Sense of Competence in Mathematics	0.90	0.80	0.91	0.81
N = 1,173				

SOURCE: 21st CCLC Evaluation, Grades 4–12 Mindsets and Behaviors Survey, 2013–14.

Appendix I. Exploring Academic Mindsets and Behaviors—Comparing SPP and ACE-Only Students at Baseline

As described in Chapter 5, the evaluation team looked at baseline differences between SPP and ACE-only students in both years of the evaluation in order to understand the starting mindsets and behaviors of the two different types of students before looking at how much they changed on their responses on the survey by the end of the school year.[40,41] Because the SPP students were recruited based on their academic needs, one hypothesis was that their mindsets and behaviors related to academic performance would be different from students who participated exclusively in ACE-only programming.

As with the staff survey described in Chapter 3, for each set of items on the survey, the evaluation team conducted Rasch analytic techniques to create scale scores for each construct. As noted earlier, a scale score places the average responses from a survey on a scale from 0 to 100, which allows evaluators to compare one set of responses with another (e.g., responses from SPP students and ACE-only students) and understand the overall response categories for each group. In this case, higher scores indicated a higher likelihood of responding that specific items were *sort of true* or *very true* (see Appendix D for more information on Rasch analysis).

Grades K–3

***Effort and Persistence* and *Learner Behavior*.** As noted earlier, for students in Grades K–3, activity leaders completed a report on mindsets and behaviors on behalf of the students. In 2012–13, that report included two key constructs, or sets of questions: *Effort and Persistence* and *Learner Behaviors*. *Effort and Persistence* items address students' perception of their abilities, willingness to work hard, and ability to focus, whereas the items in the *Learner Behaviors* construct address specific, more easily

[40] Year 2012–13 was somewhat different from 2013–14 in that, during 2012–13, the ACE-only students had been involved in the ACE program since the fall 2012 semester, whereas many SPP students had begun receiving SPP programming much later in the school year during the spring 2013 semester. Although the survey was administered to SPP students (or the activity leaders who answered on behalf of the students) at the beginning of SPP programming, this was not the case for ACE-only students in 2012–13; ACE-only students were administered the survey in the middle of the programming year. SPP students at some centers participated in ACE activities in fall 2012, but many participated in ACE activities only after the SPP programs already had been started. All of this makes comparisons between the years difficult at best.

[41] In 2012–13, only SPP and ACE-only participants who attended at least *10 hours* of SPP or ACE-only program activities were selected for inclusion in the analyses for which results are described in this section of the report. In the second year of this evaluation, the analyses included students who had participated in ACE-only programs for a year, with a minimum of 30 days of program participation. However, because the SPP program had been in operation for less than a semester in 2012–13, a 10-hour minimum threshold was determined to be a more viable cut point for identifying a regular SPP student participant. In 2013–14, because there was a full year's worth of participation between preanalysis and postanalysis, the evaluation team reverted to including in the analyses students with 30 days or more of program participation.

observed student behaviors. In 2013–14, a third set of items related to *Academic Performance* was included on the survey. These items asked school-day teachers (rather than activity leaders) to rate students on their performance in mathematics and reading. For detail on the items that make up each of the three constructs, see Appendix E, Tables E-7 and E-8.

Table I-1 shows the 2012–13 and 2013–14 scale scores for Grade K–3 students on the two constructs, *Effort and Persistence* and *Learner Behaviors*, as well as the 2013–14 scale scores for the new construct *Academic Performance*. At the onset of SPP program implementation (2012–13), ACE-only students had higher average scores on the *Effort and Persistence* construct than did SPP students, and the difference was statistically significant. This finding indicates that ACE-only students were demonstrating a slightly higher ability to do things like work hard in the program, persist even when school tasks are hard, and focus on schoolwork. It is worth noting, however, that each group of students was in the *sort of true* response range, which means that although ACE-only students were demonstrating higher levels of skill in these areas, the difference was moderate.[42] During 2013–14, there was no statistically significant difference between the SPP group and the ACE-only group in terms of *Effort and Persistence*. One reason the difference may have disappeared in 2013–14 is because, as noted earlier, in 2012–13, ACE-only students were taking the survey in the middle of their program year, and SPP students had just begun programming. In 2013–14, students of both types were taking the survey at the beginning of the program year.

A similar pattern holds for items related to *Learner Behaviors*. In 2012–13, a large difference was found between the two groups of students on the *Learner Behaviors* construct, indicating that ACE-only students were exhibiting more behaviors related to learning, such as showing interest and focus in class, seeking help from others, and working independently. In 2013–14, there was no statistically significant difference between the two groups on *Learner Behaviors*. Although it is useful to look at each year by itself to understand differences between the two types of students, differences should not be compared from one year to the next. Different groups of students were surveyed in 2012–13 and 2013–14, the surveys were administered at different times, and the survey items changed slightly.

Finally, in 2013–14, there was a statistically significant difference between the two groups of students on the *Academic Performance* construct, with ACE-only students having a higher average score and falling into the *satisfactory* (at grade-level standards) category compared with SPP students, who fell into the *needs improvement* (somewhat below grade-level standards) response category.

[42] This may be because Rasch analysis places respondents on a linear logit scale corresponding to overall response categories—that is, it is possible for there to be a statistically significant difference between two groups' mean scale scores on the logit scale while still both falling within the same overall category linked to the scale (in this case, *sort of true* on the survey's "trueness" scale). An example may clarify the point. Let us say in a study sample, the mean height of Group 1 is 6'6", and the mean height of Group 2 is 7'1". There is, presumably, a strong statistical difference between these two groups, but both would be considered *tall*. The difference is in the type of scale being used (linear logit versus categorical).

Table I-1. Grades K–3: Scale Scores on the Activity Leader *Effort and Persistence, Learner Behaviors,* and *Academic Competence* Constructs (2012–13 and 2013–14)

	Construct	Mean SPP Scale Score	Mean ACE-Only Scale Score
2012–13	*Effort and Persistence**	48.23 (Sort of True) (*n* = 471)	50.69 (Sort of True) (*n* = 330)
	*Learner Behaviors**	47.94 (Sort of True) (*n* = 471)	51.15 (Very True) (*n* = 330)
2013–14	*Effort and Persistence*	49.15 (Sort of True) (*n* = 382)	48.26 (Sort of True) (*n* = 269)
	Learner Behaviors	48.86 (Sort of True) (*n* = 382)	48.20 (Sort of True) (*n* = 269)
	*Academic Performance****	45.36 (Needs Improvement) (*n* = 365)	51.44 (Satisfactory) (*n* = 298)
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.			

SOURCE: 21st CCLC Evaluation Activity Leader Reports, 2012–13 and 2013–14.

Given the population of students the SPP programs set out to serve (i.e., those at risk for academic failure), it is not surprising to find that students in the SPP program are performing at a lower rate than are their ACE-only peers on grade-level standards or that they would be rated at a lower level on effort and learning behaviors by activity leaders in the program. We might expect SPP students to be lower on these constructs initially and for these differences to mediate across time if the SPP program is having an impact. The analysis planned for 2014–15 will examine this exact issue to understand how the program helps to improve outcomes for SPP students.

Grades 4–12

In contrast to the students in Grades K–3, students in Grades 4–12 completed the Mindsets and Behaviors Survey themselves. The survey for these students also included *Effort and Persistence* and *Learner Behavior* constructs, although the specific items were somewhat different from than those on the activity leader reports for younger students. The survey also included several other constructs that were not on the Grade K–3 activity leader reports.[43] In 2012–13, the survey included two related constructs—*Mastery Orientation* and *Self-Efficacy*—that attempted to measure students' perceptions of their own academic abilities as well as their desire to learn new things. In 2013–14, the analysis of the survey functioning, described earlier in the report, led the

[43] The survey items on these constructs referred to a level of behavior and perception that younger students would not yet exhibit in a consistent way; hence, they were not included on the survey.

evaluation team to replace the *Mastery Orientation* and *Self-Efficacy* items with a series of items related to a student's *Sense of Competence,* including the student's *Sense of Competence as a Learner* in general, and in reading and mathematics specifically (see Appendix E, Tables E-9 through E-12 for specific items). Findings from the analysis of these items are described in the following section.

***Effort and Persistence* and *Learner Behavior*.** As shown in Table I-2, in 2012–13, for students in Grades 4–12, ACE-only students scored slightly higher on both the *Effort and Persistence* and *Learner Behavior* constructs, and in both cases the difference in average scores was statistically significant. This finding indicates that ACE-only students in the older grades, similar to the younger students, are self-reporting a higher level of effort, persistence, and learning-centered behaviors than are their SPP peers.

The 2013–14 findings were somewhat similar. In that year, ACE-only students also had higher average scores on the two constructs than did SPP students. The difference in scores was only marginally significant ($p < 0.10$) for *Effort and Persistence* but was similarly significant for *Learner Behaviors.* This finding again indicates that ACE-only students may have a greater ability, at the beginning of the program, to persist, put in effort, and engage in behaviors that support learning than do SPP students.

Table I-2. Grades 4–12: Baseline Scale Scores on the *Effort and Persistence* and *Learner Behavior* Constructs (2012–13 and 2013–14)

	Construct	Mean SPP Scale Score	Mean ACE-Only Scale Score
2012–13	*Effort and Persistence**	49.85 (A Lot Like Me) (*n* = 540)	50.46 (A Lot Like Me) (*n* = 601)
	*Learner Behaviors**	49.83 (Sort of Like Me) (*n* = 540)	50.31 (Sort of Like Me) (*n* = 601)
2013–14	*Effort and Persistence**	50.87 (Sort of Like Me) (*n* = 507)	52.05 (A Lot Like Me) (*n* = 390)
	*Learner Behaviors***	50.64 (Sort of Like Me) (*n* = 507)	52.07 (A Lot Like Me) (*n* = 390)
*$p < 0.10$, **$p < 0.05$, ***$p < .01$.			

SOURCE: 21st CCLC Evaluation Youth Mindsets and Behaviors Survey, 2012–13 and 2013–14.

As with the Grade K–3 results, it is not terribly surprising that ACE-only students self-report as being slightly more competent at specific *Learner Behaviors* and *Effort and Persistence*. SPP programs have set out to recruit students who are at risk for academic failure and who therefore may be less confident or less likely to see themselves as able as their ACE-only peers to put forth strong effort and consistently focused academic behaviors.

***Self-Efficacy* and *Mastery Orientation*.** As noted earlier, the two constructs, *Self-Efficacy* and *Mastery Orientation*, were included in the Grade 4–12 survey in 2012–13 and attempted to measure student beliefs about what they can do, particularly with respect to hard work as well as actual behaviors and effort.

As shown in Table I-3, ACE-only students reported higher levels of *Self-Efficacy* than did SPP students, and the difference was statistically significant. The difference was small, and both groups indicated that the questions or items on the survey represented characteristics that were *sort of like me*, so the difference is not particularly notable. For the *Mastery Orientation* construct, there was only a slightly higher score for ACE-only students and no significant difference. However, both groups of students indicated that the questions presented characteristics that were *a lot like me,* meaning that both groups of students feel confident in their level of effort and willingness to learn new things.

Table I-3. Grades 4–12: Baseline Scores on the *Self-Efficacy* and *Mastery Orientation* Constructs (2012–13)

	Construct	Mean SPP Scale Score	Mean ACE-Only Scale Score
2012–13	*Self-Efficacy**	48.46 (Sort of Like Me) (*n* = 547)	50.11 (Sort of Like Me) (*n* = 605)
	Mastery Orientation	50.21 (A Lot Like Me) (*n* = 547)	50.73 (A Lot Like Me) (*n* = 605)
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.			

SOURCE: 21st CCLC Evaluation Youth Mindsets and Behaviors Survey, 2012–13 and 2013–14.

Together, these findings indicate that, although ACE-only students self-reported slightly higher levels of skill on *Self-Efficacy* and *Mastery Orientation* than did SPP students, the differences are so minor that they should be interpreted with caution. The more important story here is that the scores suggest both groups of students viewed their school effort in a positive way—thus, the high score on *Mastery Orientation*—and had some doubt about their willingness to work hard or the value of hard work—hence the score on *Self-Efficacy*. Two possible explanations exist for the high scores (with the majority of students choosing the highest response category) and lack of variation between the two types of students: (1) the poor functioning of the survey—as noted earlier, these two scales were removed in the 2013–14 survey, or (2) that students are more likely to believe that their academic skills and capabilities are inherent and not able to be changed through hard work. Focusing on improving students' sense of *Self-Efficacy* may be an important step for programming in the future.

***Sense of Competence*.** In 2013–14, three new constructs were added to the survey to replace *Mastery Orientation* and *Self-Efficacy*. These new constructs, *Sense of Competence as a Learner*, *Sense of Competence as a Reader*, and *Sense of*

Competence in Mathematics, were designed to understand how well students perceive their abilities to perform well on school work overall as well as on specific subjects.

As Table I-4 shows, there was no significant difference between the responses of SPP students and ACE-only students on any of the three constructs. ACE-only students reported slightly higher levels of competence in all three areas, but none of those differences were statistically significant, and in all three areas for both types of students, they responded that the characteristics described in the survey were *sort of like me.* This finding is somewhat surprising, given the fact that findings from other constructs on the survey showed that ACE-only students reported higher levels of functioning on almost all the constructs (see Tables I-1 through I-3) and the fact that SPP students are theoretically more at risk academically than are their ACE-only peers. The evaluation team expected to see a greater difference between students on these *Sense of Competence* constructs. Further analysis in 2014–15 will be important for exploring these differences.

Table I-4. Grades 4–12: Baseline Scores on *Sense of Competence* Constructs (2013–14)

	Construct	Mean SPP Scale Score	Mean ACE-Only Scale Score
2013–14	*Sense of Competence as a Learner*	50.90 (Sort of Like Me) (*n* = 508)	51.19 (Sort of Like Me) (*n* = 391)
	Sense of Competence as a Reader	50.22 (Sort of Like Me) (*n* = 508)	51.07 (Sort of Like Me) (*n* = 391)
	Sense of Competence in Mathematics	51.04 (Sort of Like Me) (*n* = 507)	51.25 (Sort of Like Me) (*n* = 390)
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$ – in this case there were no significant findings.			

SOURCE: 21st CCLC Evaluation Youth Mindsets and Behaviors Survey, 2012–13 and 2013–14.

Summary of Youth Mindsets and Behaviors at Baseline

Overall, the survey data for both types of students and across the constructs suggests that there were some differences between the ACE-only and SPP students in their academic mindsets and behaviors at the time programming began, both in 2012–13 and 2013–14. These results may indicate that SPP students have certain academic-related mindsets and behaviors to a lesser extent than do ACE-only students. As noted earlier, this may be because SPP programs target students at risk for academic failure (and who are, therefore, potentially less academically capable or confident than are ACE-only students). On the other hand, limitations to the survey and survey administration also may explain the differences between the two groups of students. As noted earlier, many ACE-only students in the 2012–13 evaluation had participated in OST activities longer than the SPP students had, which means they may have had more exposure to

tutoring, homework help, and other enrichment programming that supports the development of academic mindsets and behaviors. This does not explain the 2013–14 results, however, because in that year students took the survey at the same time, indicating that survey administration cannot account fully for observed differences.

Although the exact reason for differences between SPP and ACE-only students at baseline is unknown, the value of this analysis was to uncover what differences may exist and highlight areas for future study in the 2014–15 analysis. In particular, future study will focus on the pre-post differences in SPP students compared with those in ACE-only students and whether the intervention appears to minimize some of the differences found between the two types of students at the program's outset.

Appendix J. High School Pilot Survey Protocol

Protocol for High School Survey Pilot

General Instructions for Site Coordinators:

Please identify 25 students in your out of school ACE programming to take the survey. The students should be in Grades 9–12 and participate in ACE programming. They may be students who have already taken a student survey as part of the 21st CCLC evaluation, but they do not have to be.

Please send a consent form home with the student and wait six calendar days before surveying the student. If the parent declines, please choose another student to participate.

You can survey the students in one large group or small groups. Please try to complete all 25 surveys by the end of May and mail them along with the attached template completed with your notes in the provided envelope to the attention of:

Nicole Adams
American Institutes for Research
1120 E. Diehl Road, Suite 200
Naperville, IL 60563

If you have any questions, please contact Nicole Adams at tx21stCCLC@air.org or (630) 649-6652.

How to Administer the Survey

Hand each student a survey. They should not put their name on the survey. Before they begin, tell the students the following:

> "The purpose of this survey is to find out more about students who participate in the TX 21st Century Community Learning Centers (CCLC) out of school time program so the state of Texas can keep making programs better for you and other students at your school. Because we would like to include you in this study, you are being asked to participate in a survey. It should take approximately 15 minutes for you to complete. Below are questions that ask about you and some of the things you think about and feel about yourself at school and as a learner. This is not a test. There are no "wrong" answers. Please choose the answer that is most true for you.
>
> Your parents have been informed about your opportunity to participate. Still, this survey is completely voluntary. You do not have to answer any of the questions if you don't want to, and you can stop doing this survey at any time. If you do not want to take the survey, turn your paper over and find something to do quietly.
>
> Do not put your name on the survey. Everything you write is confidential and anonymous to the extent permitted by law, which means that no one (not your parents, teachers, school staff, or other students) will be allowed to know how you answer these questions. If you have any questions about the survey let me know and I can give you a phone number you can call.
>
> When you are done with your survey, turn it over and wait quietly. When everyone is done, I am going to ask you a few questions about what you thought of it. The people who made the survey want to make sure it is easy to understand so that it can be used next year with all of the high school students participating in out of school time programs like this one."

If anyone does not want to participate you can just collect a blank survey from them. They do not have to do the survey. If anyone has questions, answer any questions you can, but refer them to the phone number on the survey itself. That is a number for someone at AIR who can answer their questions.

After everyone has finished the survey, ask them the following three questions. If it is a large group, please make sure you are encouraging everyone to answer. Try to get as many responses as possible to each question without pressuring students to respond if they do not want. If you get short answers, probe the students more. Try to get as much information as you can. Record the answers on the attached page and send it along with the completed surveys in the envelope provided.

Which questions were confusing or hard to answer? Why?

What did you think of the answer choices? Was it hard or easy to pick your answer? Why?

What would you change about the survey to make it better for high school students like yourselves?

Questions for Survey Administrator

Questions for survey administrators to answer – these ask your opinion about the survey	Notes
How long did the survey take?	Record the shortest completion time: Record the longest completion time:
Did students ask you any questions? What were they?	
What do you think of the questions? Are there any you think are confusing or hard to understand?	
What do you think of the answer choices? Too many? Too few?	
Any other feedback about the survey?	

Questions for Students

Question	Notes
Ask students the following questions	
Which questions were confusing or hard to answer? Why?	
What did you think of the answer choices? Was it hard or easy to pick your answer? Why?	
What would you change about the survey to make it better for high school students like yourselves?	

Appendix K. High School Pilot Survey Results

High School 1

Student Skills and Beliefs High School Survey Pilot Results
25 students completed the pilot survey

1. When you do not come to this ACE program, what are you usually doing instead? (check one.)

I go home.	29%
I have to take care of younger brothers, sisters, other children, or other family members.	18%
I go to a friend's house.	12%
I go to the home of an adult neighbor, relative, or family friend.	0%
I go to organized non-academic after-school activities (sports teams, clubs, etc.).	0%
I go to organized academic after-school activities (e.g., tutoring programs).	0%
I go to private lessons or classes (music, sports, dance, etc.).	6%
I have a job.	23%
Some other place	12% (Church, sleeping)

	Mean (where 1 is not at all like me and 4 is exactly like me)
Doing well in school is an important part of who I am.	2.64
Getting good grades is one of my main goals.	2.64
I am the kind of person who takes pride in doing my best in school.	2.52
Getting a college education is a high priority for me.	3.24
The things I am learning in school will be important later in life.	2.72
I think most of my classes are interesting.	2.12
I like to learn new things.	2.60
I need to finish high school to get a good job.	3.04
The things I am learning in school will be useful in a job or career.	2.79
My intelligence is something that I can change if I want to.	2.84
I can always solve difficult problems if I try hard enough.	2.64
I can always do a good job if I try hard enough.	3.00
It is important to me to learn as much as I can.	2.80
I am a hard worker.	3.12

	Mean (where 1 is not at all like me and 4 is exactly like me)
I finish whatever I begin.	2.84
I continue steadily toward my goals.	2.92
I don't give up easily.	2.75
I try hard in school.	2.67
When I don't understand my homework, I get help so I can get it done.	2.48
I set goals for myself. For instance, things I want to learn or get better at.	2.68
I feel proud to be part of my school.	2.60
My teachers take the time to get to know me.	2.12
It's very easy for me to get along with other teens.	2.52
When I meet someone new, I know he or she will like me.	2.56
I get along with friends as well as other teens my age.	2.88
It's easy for me to join a new group of teens.	2.60

2. Thinking ahead to your future, what do you think will happen?

	Mean (where 1 is probably won't and 3 is definitely will)
I will get good grades and be successful in high school.	2.5
I will graduate from high school.	2.7
I will go to college.	2.4

3. Below are some statements that might describe how you feel about your ACE program. For each statement, please indicate how true the statement is for you.

	Mean (where 1 is not at all true and 4 is completely true)
I like coming here.	2.46
I have fun when I'm here.	2.54
I can always find things that I like to do here.	2.17
I learn new things here.	2.67
I explore new ideas.	2.42
I get to do things that I have never done before.	2.38
I feel like I belong here.	2.42
I feel like my ideas count here.	2.35
I feel like I am successful here.	2.58
I feel comfortable at this program.	2.58
I like the adults in this program.	2.71

Please tell us a little bit more about how you feel about this ACE program:

4. Why did you first choose to come to the program?

   Friends/knew people (*N*=3)

   Extra help/get better in school (*N*=6)

   Wanted to learn something new (*N*=3)

   To stay out of trouble (*N*=1)

   No choice (*N*=2)

5. What is the #1 reason you stay in the program?

   To keep learning/get help (*N*=7)

   It's fun (*N*=3)

   Program Director (*N*=2)

   Free stuff (*N*=1)

   Boredom (*N*=1)

   I don't know (*N*=2)

6. What do you like most about the program?

   Learning something new/extra help (*N*=4)

   It's fun (*N*=3)

   Teachers/Program Director/the people (*N*=6)

   Free stuff/snacks/candy (*N*=3)

   Don't know (*N*=2)

High Schools 2 and 3 Combined

Student Skills and Beliefs High School Survey Pilot Results
53 students completed the pilot survey

1. When you do not come to this ACE program, what are you usually doing instead? (check one.)

I go home.	46%
I have to take care of younger brothers, sisters, other children, or other family members.	10%
I go to a friend's house.	12%
I go to the home of an adult neighbor, relative, or family friend.	0%
I go to organized non-academic after-school activities (sports teams, clubs, etc.).	9%
I go to organized academic after-school activities (e.g., tutoring programs).	3%
I go to private lessons or classes (music, sports, dance, etc.).	3%
I have a job.	10%
Some other place	7% (ROTC, church, sleeping)

	Mean (where 1 is not at all like me and 4 is exactly like me)
Doing well in school is an important part of who I am.	3.02
Getting good grades is one of my main goals.	3.11
I am the kind of person who takes pride in doing my best in school.	2.87
Getting a college education is a high priority for me.	3.34
The things I am learning in school will be important later in life.	2.94
I think most of my classes are interesting.	2.60
I like to learn new things.	3.40
I need to finish high school to get a good job.	3.47
The things I am learning in school will be useful in a job or career.	3.02
My intelligence is something that I can change if I want to.	3.12
I can always solve difficult problems if I try hard enough.	3.34
I can always do a good job if I try hard enough.	3.45
It is important to me to learn as much as I can.	3.29

	Mean (where 1 is not at all like me and 4 is exactly like me)
I am a hard worker.	3.26
I finish whatever I begin.	2.96
I continue steadily toward my goals.	3.08
I don't give up easily.	3.34
I try hard in school.	2.98
When I don't understand my homework, I get help so I can get it done.	2.75
I set goals for myself. For instance, things I want to learn or get better at.	3.12
I feel proud to be part of my school.	2.98
My teachers take the time to get to know me.	2.62
It's very easy for me to get along with other teens.	3.38
When I meet someone new, I know he or she will like me.	2.98
I get along with friends as well as other teens my age.	3.35
It's easy for me to join a new group of teens.	3.06

2. Thinking ahead to your future, what do you think will happen?

	Mean (where 1 is probably won't and 3 is definitely will)
I will get good grades and be successful in high school.	2.31
I will graduate from high school.	2.40
I will go to college.	2.19

3. Below are some statements that might describe how you feel about your ACE program. For each statement, please indicate how true the statement is for you.

	Mean (where 1 is not at all true and 4 is completely true)
I like coming here.	3.04
I have fun when I'm here.	3.02
I can always find things that I like to do here.	3.02
I learn new things here.	3.06
I explore new ideas.	2.98
I get to do things that I have never done before.	2.80
I feel like I belong here.	2.79
I feel like my ideas count here.	2.65

	Mean (where 1 is not at all true and 4 is completely true)
I feel like I am successful here.	2.92
I feel comfortable at this program.	3.13
I like the adults in this program.	3.17

Please tell us a little bit more about how you feel about this ACE program:

4. Why did you first choose to come to the program?

   Wanted to learn something new/clubs and activities (*N*=15)

   Extra help/get better in school (*N*=9)

   Friends/knew people/meet new people (*N*=8)

   Don't want to go home/stay out of trouble/boredom (*N*=7)

   Snacks (*N*=6)

   Program environment (*N*=2)

   Freshmen year/sophomore year (*N*=2)

   I don't know (*N*=1)

5. What is the #1 reason you stay in the program?

   Activities/overall program experience (*N*=13)

   Friends/people in program (*N*=10)

   Continue learning/help with school work (*N*=9)

   Snacks (*N*=9)

   Boredom/nothing else to do/Don't want to go home (*N*=1)

   I don't know (*N*=1)

6. What do you like most about the program?

   Learning something new/extra help/activities (*N*=15)

   It's fun/program environment (*N*=7)

   Friends/the people (*N*=9)

   Snacks (*N*=8)

   Everything/don't know (*N*=6)

   Keeps me out of trouble/nothing else to do (*N*=2)

   Get to relax (*N*=2)