Texas 21st Century Community Learning Centers

Year 2 Evaluation Report

AIR®
AMERICAN INSTITUTES FOR RESEARCH®

Prepared for: Texas Education Agency

Table of Contents

List of Acronyms

21st CCLC	21st Century Community Learning Center
ACE	Afterschool Centers on Education
AIR	American Institutes for Research
AEIS	Academic Excellence Indicator System
APT-O	Assessment of Afterschool Practices Observation Tool
CSF	Critical Success Factor
CSM	Critical Success Model
ESEA	Elementary and Secondary Education Act
HLM	Hierarchical Linear Modeling
MRFM	Multi Facet Rasch Measurement
NCLB	No Child Left Behind
OCE	Observation of Child Engagement
PAMS	Profile Analysis via Multi-dimensional Scaling
PEIMS	Public Education Information Management System
POS	Point of Service
PPICS	Profile and Performance Information Collection System
PQA	Program Quality Assessment
SEA	State Education Agency
TAKS	Texas Assessment of Knowledge and Skills
TEA	Texas Education Agency
TX21st	TX21st CCLC Student Tracking System

[intentional blank page]

Texas 21st Century Community Learning Centers

Year 2 Evaluation Report

February 2013

Prepared for:

Texas Education Agency

AIR

AMERICAN INSTITUTES FOR RESEARCH®

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

CREDITS

Contributing Authors

American Institutes for Research
Neil Naftzger
David Manzeske, Ph.D.
Mary Nistler
Andrew Swanlund

Gibson Consulting Group, Inc.
Amie Rapaport, Ph.D.
Joseph Shields

David P. Weikart Center for Youth Program Quality
Charles Smith, Ph.D.
Anna Gersh
Samantha Sugar

Submitted to

[intentional blank page]

Executive Summary

The 21st Century Community Learning Centers (21st CCLC) program, authorized under Title IV, Part B, of the Elementary and Secondary Education Act (ESEA), as amended by the No Child Left Behind Act of 2001 (NCLB), supports the creation of community learning centers that provide academic enrichment opportunities during non-school hours for children, particularly students who attend high-poverty and/or low-performing schools.[1] The federal grants are awarded to state education agencies (SEAs), which, in turn, make competitive awards to eligible grantees to support afterschool and summer learning programs.[2] In July 2002, the federal government awarded the Texas Education Agency (TEA) $24.5 million to fund TEA's first cohort of 21st CCLC grantees for the 2003–04 school year. As of 2010–11, this and subsequent federal funding has resulted in 281 grants being awarded in Texas over seven 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.[5] These five areas were examined by the evaluation team, and findings on gains made toward meeting the performance objectives will be presented in the Year 2 Evaluation Report.

To ensure that grantees funded by the ACE program are positioned to achieve program objectives, TEA has developed a research-based Critical Success Model (CSM). This model includes four Critical Success Factors (CSFs) which represent behavioral changes that should be demonstrated by students and families enrolled in the program, or by the adults working on their behalf, to ensure success in meeting programmatic goals and objectives. The CSFs and corresponding milestones (key strategies that establish the foundation on which critical success factors are built) follow.[6]

---

[1] For more information see http://www2.ed.gov/policy/elsec/leg/esea02/pg55.html

[2] Grantees include local education agencies, non-profits, 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. While 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.

[5] 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

[6] Beginning in 2009–10, with the sixth funding cycle, program guidelines were revised to require grantees to develop and implement programs in alignment with TEA's research-based CSM. Cycle 6 grantees were required to use this model to establish program goals and implement their programs. In addition, Cycle 6 grantees must collect and report performance measure data to TEA based on milestones and CSFs. Cycle 5 grantees are not subject to these requirements, although they were made aware of the

- CSF1 emphasizes both student and family engagement. Student engagement and family engagement are not necessarily achieved together, because they require different strategies and activities. The milestone for this CSF is predicated on the implementation of research-based, innovative instructional techniques and opportunities that encourage student and family engagement.

- CSF2 addresses student involvement in school, exemplified through more participation in extracurricular activities, and more mentors supporting students. The milestone for this CSF emphasizes the role of adults as advocates for students.

- CSF3 addresses the use of assessment data to evaluate and revise student activities and services. A milestone strategy is for program staff to conduct ongoing and continuous assessments to identify student needs and how to revise program services.

- CSF4 addresses staff professional development. The milestone strategy is for programs to provide all staff the "required training opportunities," which are then implemented in the afterschool program.

Overview of the Evaluation

Through a contract with TEA, the evaluation is being conducted by the American Institutes for Research (AIR) and its partners, Gibson Consulting Group, Inc. (Gibson Consulting) and the David P. Weikart Center for Youth Program Quality (the Weikart Center), from January 2011 through August 2013 and possibly through August 2015 if additional funding is identified. To date, state evaluation efforts have been focused on programs that were awarded grants during Cycle 5 and Cycle 6.[7] The overarching goal of the evaluation is to determine which program strategies and approaches are most effective within particular contexts in encouraging student behaviors (CSFs) that lead to improved student outcomes. As such, the evaluation is designed to address two primary research objectives:

- **Research Objective 1:** To conduct a statewide assessment of ACE programs, operations, participation, and student achievement;

CSM when Cycle 6 was first implemented and are encouraged to use the model as a guide for improving programs.
[7] At the beginning of the 2010–11 school year, Cycle 5 programs began their third year of implementation, and Cycle 6 programs began their second year of implementation. Programs funded by the Texas 21st CCLC program are funded for five years. Applicants that received funding in previous 21st CCLC cycles are eligible to apply for funding to serve new or existing programs; however, the pre-existing grant must expire before the project start date of the new grant cycle.

- **Research Objective 2:** To identify and describe innovative strategies and approaches implemented by successful Cycle 5 and Cycle 6 grant-funded programs.[8]

The second year of the evaluation extends the work conducted in the first year of the evaluation, for which findings were presented in the Interim Report. As in the first year, data collection and analysis in the second year of the evaluation focused on program quality—how it was demonstrated in program activities and organizational processes, and its impact on student participation and student academic and behavioral outcomes. By the second year, observation and interview data had been collected on 80 ACE programs (in spring and fall 2011). These data were used first to select 15 centers that were visited again in spring 2012, with selection of those centers based on indicators of quality. The data from the 2011 site visits were also used to examine more deeply the instructional practices in sessions where students were consistently engaged. At the 15 centers visited in 2012, comprehensive descriptions of program activities were developed from observations, and summaries of organizational practices were developed from interviews with key stakeholders (the site coordinator, principal, and project director, as well as teachers whose classes were observed).

The evaluation in the second year included extensive analysis of staff survey data (collected in 2011 at the centers where site visits were conducted), and observation data, also collected in 2011, which included three observation protocols (the Youth Program Quality Assessment, PQA; the Observation of Child Engagement protocol, OCE; and portions of the Assessment of Afterschool Practices Observation Tool, APT-O) that address academic content. For each of the protocols, observers applied scores on the included constructs. The data available from the TEA provided information on program and student characteristics, available through the TX21st CCLC Student Tracking System (TX21st); as well as student outcome data, using Texas Assessment of Knowledge and Skills (TAKS) scale scores in reading and mathematics as outcomes. Additional information on student demographics, discipline incidents, school-day attendance, and grade promotion came from the Public Education Information Management System (PEIMS). Campus-level performance data from the Academic Excellence Indicator System (AEIS) was used to create a non-participant comparison group.

---

[8] These research objectives may be expanded to include new grantees in later years of the evaluation.

Overview of Findings

The issue of program quality can be conceptualized as the binding thread that is woven throughout the CSM, adopted by TEA to guide implementation of ACE programming in a way that is likely to impact the lives of participating youth in meaningful and lasting ways. The 2010–12 evaluation was guided by seven research questions, most of which address quality practice in ACE programs, and how program quality influences student outcomes. Each of these questions is outlined below, along with a summary of key findings resulting from analyses oriented at answering each question.

1. *What instructional approaches are associated with high levels of student engagement at the point of service?*

The analysis of observation data from both the 2011 and the spring 2012 site visits showed that three instructional approaches distinguished high quality activities. One of these was clarity of purpose, whereby the activities were clearly designed to achieve explicit objectives. In the high quality academic enrichment sessions, the instructors clearly stated the learning objectives and then led students through a variety of learning activities related to those objectives. In the non-academic enrichment sessions, the objectives may not have been as explicitly stated, but the activities were still purposed toward learning within the context of the sessions.

A second feature was intentional use of time. This feature, which was anchored in planning and pacing, was found to be essential for keeping students busy and engaged throughout the observed sessions. Materials were ready when the sessions began. Routines were worked into sessions so little time was wasted when students began sessions and transitioned from one activity to another. The pace was generally quick, and, as a result, student accomplishments by the end of the activities were evident.

The third feature was an active and interactive instructor, who continually engaged with students, even when students were working in small groups or on their own. The instructors moved about the room, looked over students' shoulders, asked questions that deepened student knowledge, noticed and helped when students had trouble, and managed student behaviors before any students became disruptive.

2. *What organizational processes are found to be drivers of instructional/point of service quality at high performing centers?*

Several organizational processes were examined to determine whether they were drivers of instructional and point of service quality. Center intentionality was a key dimension that was examined. Center intentionality refers to the center's purpose and the degree to which the purpose is defined and appropriate resources provided. Among the 15 ACE programs in the sample, those with high center intentionality with respect to academic enrichment programming showed a strong focus on meeting academic objectives that were emphasized during the school day core classes. This was

particularly evident in the elementary centers and one middle school center that had developed curricula for their academic enrichment activities and consistently aligned the curricula to school-day learning objectives. Intentionality with respect to non-academic enrichment programming was uneven across all grade levels. Across the centers, there did not appear to be a shared understanding of what non-academic enrichment programming is. Numerous respondents referred to these activities as opportunities to relax, play games, and be physically active—thus, they were non-academic but did not provide opportunities to learn new knowledge and skills associated with a non-academic subject area or discipline.

Practices to monitor for improvement (that is, to improve the quality of activities) were evident in the majority of the centers in the site visit sample. Particularly strong were the centers that systematically approached monitoring of academic enrichment activities and modifying curricula and instructional approaches to better engage students and meet learning objectives. To an extent, this was more obvious among the elementary centers in the sample. In the secondary centers in the sample, where monitoring was assessed as high, attendance was typically the foremost factor in monitoring and making improvements in programming. Site coordinators and, in some cases, project directors considered student attendance and student motivation to participate—that is, were students interested in the activities? Were the sessions themselves appealing? Were there barriers to attendance that needed to be addressed? Responses to monitoring of this type included implementing different types of activities, particularly non-academic enrichment activities, based on student interest, and working with instructors to better align instructional methods with students' developmental needs. This type of monitoring seems appropriate in high schools where students have a high level of choice regarding participation.

Nearly all of the 15 centers showed a clear linkage to the school day. Close ties with the school were reflected in information transfers related to students and school learning objectives, availability of space, and administrative support.

Staff development and staff collaboration were other means of supporting program quality. Among the 15 centers in the 2012 site visit sample, staff development and opportunities for collaboration, particularly collaboration in scheduled sessions, were more available to instructors who were not certified teachers than those who were teachers during the school day. The school-day teachers had little time to meet formally, although they had some opportunities for sharing information on an informal basis. The scheduling constraints, a site coordinator said, made it difficult to improve instructional methods of staff in order to make them more engaging for youth. In this group of ACE programs, the elementary centers were more reliant on non-certified teachers (usually affiliated with the grantee organization) to instruct sessions than secondary centers, where activities were frequently instructed by certified teachers. Thus, it was elementary

centers in the sample—more than secondary centers—that provided opportunities for both staff development and collaboration for at least a majority of instructors. Providing opportunities for staff development and collaboration for certified teachers may be a persistent challenge for many afterschool programs. Sometimes, this is addressed through the professional development of site coordinators, which was extensive for the majority of the centers in the sample.

Finally, community connections were important to many of the centers, most of which had developed partnerships with organizations and agencies that resulted in expanded opportunities for programming, youth, and parents.

3. What innovative strategies and approaches can be identified from these centers that warrant replication and emulation?

Many innovative strategies and approaches from centers serving youth in elementary, middle, and high schools are presented in the Year Two Evaluation Report. They are too numerous to list in this summary chapter, but they may be referred to as models that might be replicated. It is important to note that although the type of activities that were observed and presented as exemplars are varied, they all exemplified the three instructional features described in Chapter 2: clarity of purpose, intentional use of time, and an active and interactive instructor.

4. What is the relationship between the characteristics of individual youth, center quality, and other center characteristics and levels of student participation in ACE programming?

Getting students to participate in ACE-funded programming consistently and on a sustained basis over time is a critical first step in enhancing the likelihood that students will achieve desired program outcomes. A hypothesis was that students enrolled in centers demonstrating higher quality would be more likely to participate in more total hours of programming, and for a longer duration, during the course of the 2010–11 school year (measured by the number of days between the first day of participation and the last day of participation).

In order to test this hypothesis, centers visited in 2011 were classified into different quality profile types based on observation and staff survey data collected during this period. Four profile types were defined based on observation data: 1) *high POS (point of service) quality*; 2) *low POS quality*; 3) *high APT-O/Academic Climate*; and 4) *high OCE*. It was expected that participation-related outcomes would be better in centers classified in the *high POS quality* cluster, because centers assigned to this cluster were characterized by high program quality on each of the three observation protocols employed: the PQA, the OCE, and the APT-O. In addition, centers assigned to the *low POS quality* cluster were expected to do less well on participation-related outcomes

because of lower levels of observed program quality, based on the same three protocols.

When multilevel models were run to explore these relationships, the hypothesized pattern of results was found: students enrolled in *high POS quality* centers participated in programming for a significantly longer duration during the 2010–11 school year, while students enrolled in *low POS quality* centers participated for a significantly shorter duration.

However, a similar result was not found when the total hours of ACE programming attendance was used as an outcome. Neither *high POS quality* nor *low POS quality* was predictive of this outcome.

Surveys completed by activity leaders asked respondents to report the extent to which they engaged in quality-related practices. From the analysis of these responses, two quality-related clusters were identified: (1) *lower reported quality* centers, and (2) *higher reported quality* centers. The hypothesis was that centers in the *higher reported quality* cluster would demonstrate better participation-related outcomes than centers enrolled in the *lower reported quality* cluster. When the total hours of participation in ACE programs was used as an outcome, the hypothesized relationships were found, suggesting that the implementation of higher quality practices was related to more hours of participation. However, implementation of these same practices was not found to be related to the duration of participation.

Overall, the approach adopted by the evaluation team to construct quality profiles based on observation and staff survey data did not yield variables that were consistently predictive of both the total number of hours of ACE participation and the duration of participation. However, each analysis resulted in significant findings that supported, at some level, the hypothesized relationship between program quality and participation outcomes.

5. Does the impact on student outcomes vary by relevant ACE program characteristics, including center quality?

One of the primary objectives of the ACE evaluation is to understand the relationship between participation in ACE programs and student improvement, particularly improvement on outcomes related to academic performance, school-day attendance, disciplinary incidents, and promotion rates. It is these outcomes toward which ACE programs are to direct their programming.

Using the observation and staff survey-based quality clusters employed to answer research question 4, a three-stage analytic strategy was developed to assess how program quality was related to the effect of participating in ACE programming on a variety of student outcomes associated with the 2010–11 programming period:

- TAKS-Reading/ELA and TAKS-Mathematics scores
- The number of school-day absences
- The number of disciplinary incidents
- Grade level promotion

It was hypothesized that centers demonstrating higher quality would have a stronger, positive effect on each of these outcomes, while lower quality centers would have a weaker effect. This hypothesis was borne out in the following findings:

- Centers assigned to the *high POS quality* cluster were found to have *higher* effect sizes in terms of supporting a decrease in disciplinary incidents than centers assigned to other quality types

- Centers assigned to the *high POS quality* cluster were found to have *higher* effect sizes in terms of supporting student grade promotion than centers assigned to other quality types

- Centers assigned to the *low POS quality* cluster were found to have *lower* effect sizes in terms of supporting student performance on the TAKS-Reading/ELA assessment than centers assigned to other quality types.

No relationship was found between higher quality programming and larger effect sizes in terms of program impact on TAKS-Mathematics scores and school-day absences.

6. *To what extent do students who have higher participation rates demonstrate better academic and behavioral outcomes as compared with similar students who participate in 21st CCLC at lower levels?*

Analyses were undertaken to explore the extent to which students who attended programming for 60 days or more demonstrated better outcomes than similar students who participated in ACE programming for 30 to 59 days. Results from these analyses demonstrated that higher levels of attendance in ACE-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. However, the effect of higher levels of attendance on TAKS-related outcomes was quite small. Higher levels of attendance in ACE programs proved to be more impactful in terms of reducing disciplinary incidents and school-day absences and particularly in supporting grade promotion. In the latter case, students attending 60 days or more had a rate of grade promotion 23% to 40% higher than students attending 30 to 59 days. This information provides ACE programs with additional understanding regarding how much additional program impact can be derived from keeping students engaged in ACE programming for 60 days or more.

7. To what extent do students participating in services and activities funded by 21st CCLC demonstrate better achievement (along with other student outcomes) as compared to similar students not participating in the program?

Analyses were undertaken to assess the impact of the ACE program on student outcomes by comparing ACE program participants with students who were similar in all observable ways except program attendance. Program participation was defined in two separate ways to create a sharper contrast between participants and non-participants. A group of "low" program participants was identified as having participated in at least 30 days of programming, and these students were compared to students who did not participate in any ACE programming. A group of "high" program participants was identified as having participated in at least 60 days of programming, and these students were also compared to students who did not participate in any ACE programming.

For both low- and high-attending students, ACE program participation had a statistically significant impact on TAKS scores, discipline, absences, and grade promotion for many of the grade levels relative to students that did not participate in the program.

- For Grades 9–12 only, ACE program participation was associated with higher TAKS scores in Reading/ELA and Mathematics.

- For Grades 6–12 only, ACE program participants had fewer disciplinary incidents than non-participating students.

- For low-attending students in Grades 4–5 and high-attending students in Grades 4–11, program participation was associated with fewer school-day absences.

- Program participation was associated with increased likelihood of grade promotion in Grades 6–11 for low-attending students and in Grades 4–11 for high-attending students.

The magnitude of each of these program effects was primarily in the small to moderate range, with the largest effects associated with reductions in school-day absences and grade promotion. However, for both low- and high-attending students, impacts on grade promotion were especially substantial in Grades 9–11. In this case, participation in an ACE program increased the likelihood of being promoted to the next grade level by 79% and 97%, respectively. The magnitude of this effect size is large.

In addition, while the impact of the program on each of these outcomes was found to be significant in the Interim Evaluation Report, the level of impact was significantly larger for students in both the low- and high-attending groups, demonstrating the importance of retaining students in programming beyond the 30-day threshold.

Primary Themes and Recommendations

Most of the findings outlined in this report can be distilled down to two primary themes, both of which can guide future evaluation work and inform how TEA approaches the design and delivery of training, technical assistance, and professional development for staff working in ACE-funded programs:

- *Program Quality Matters.* In some instances, measures of program quality employed during the evaluation to assess center functioning in the adoption of practices to support academic skill-building and youth development were related to both student participation in ACE programs and the achievement of student outcomes. In particular, measures of program quality were found to be related to levels and duration of participation in ACE programming, a decrease in disciplinary incidents, grade promotion, and performance on TAKS-Reading/ELA. Each of the measures employed to formulate quality estimates detail specific practices that program staff can adopt to support implementation of quality programming that supports academic skill-building and mastery and youth development among participating students. TEA should consider reviewing these measures to see how the practices articulated in each measure may further inform the formulation and delivery of training, professional development, and technical assistance oriented at supporting centers in improving the quality of their offerings. Each of the tools operationalize the features of high quality activities noted in this report, helping to ensure clarity of purpose and intentional use of time, and to provide markers for the types of behaviors that define an active and interactive instructor.

  In addition, states are increasingly working on the development and implementation of quality assessment tools and mechanisms, such as leading indicator systems, to feed data on program quality back to 21st CCLC programs to support quality improvement efforts. Most of these systems are predicated on supporting program adoption of specific quality-related practices. TEA is encouraged to review their current efforts in this regard to see what additional approaches could be implemented to get actionable quality data into the hands of program administrators and staff.

- *High School Students Especially Benefitted from ACE Participation.* Consistently, across each of the outcomes examined, program effects were found to be the greatest for high school students participating in ACE programming. (The 15 centers in the 2012 site visit sample showed that academic support was mainly available through tutorials that were highly aligned to the school day, which may be one reason for the benefits for high school students.) In some instances, these effects were quite large, particularly in relation to grade promotion, where students attending 60 days or more demonstrated a 97% better chance of being

promoted to the next grade level as opposed to similar students not enrolled in the program.

Such results warrant further examination into what attracts these older youth to ACE programming, and what keeps them participating for an extended period of time. Engaging in such an examination is of particular importance since efforts to identify the features of high quality offerings detailed in this report consistently demonstrated lower performance on the part of high school programs on key facets of program quality. This may suggest there is something unique about high school students that opt to participate in ACE programming that makes it particularly likely they will benefit from their participation in such programming. It is the hypothesis of the research team that student motivation plays a strong role in how high school students interact with and benefit from their participation in ACE programming. Understanding what role motivation plays in how high school students connect with afterschool programming would be especially helpful to TEA and the field in understanding how to reach out and engage older youth in a manner that is likely to lead to positive academic achievement outcomes.

The information in this report provides concrete evidence for, and examples of, how program quality can support the achievement of desired ACE program outcomes. What has not been measured or assessed to date is how youth change as a *direct consequence* of ACE program participation, and how these changes transfer outside the program to impact the types of student academic and behavioral outcomes examined in this report. These more immediate, *within-program outcomes* can fall within a wide spectrum of categories, including social emotional learning; critical thinking and decision-making; initiative and self-direction, and so on (Wilson-Ahlstrom, Yohalem, DuBois, & Ji, 2011). In addition, acquisition of content-specific skills in areas like reading and mathematics are likely to be more targeted in nature within a given ACE program, both in terms of the area of emphasis within a given program and the student's unique area of need, particularly in relation to students falling below proficiency. As we move into Year 3 of the evaluation, we will focus more on understanding how programming impacts more immediate student skills and functioning that translate into desirable academic and behavioral outcomes.

More specifically, with the onset of the 2012–13 school year, TEA will be providing current ACE grantees that have demonstrated a capacity to provide higher quality programming with the opportunity to obtain a State of Texas Assessments of Academic Readiness (STAAR) Supplemental Academic Support Grant to better identify and serve students who are particularly in need of help and support in developing STAAR-related skills. In many respects, the decision by the TEA to develop such a program provides a number of opportunities to structure the Year 3 21st CCLC evaluation in a way that

allows for a variety of research questions to be addressed that further build on and extend evaluation efforts conducted to date, including:

1. How are recipients of STAAR Supplemental Academic Support Grants using these funds to identify and recruit high need students into programming; what steps are being taken to align programming with the CSM adopted by TEA for the 21st CCLC program; and to what extent is programming being delivered in a manner that is consistent with afterschool quality frameworks?

2. What characteristics are associated with Supplemental Academic Support Grant-supported activities where there are high levels of youth-reported engagement?

3. What impact does programming funded by the STAAR Supplemental Academic Support Grants have on short-term program outcomes, like student task persistence, motivation, and academic self-efficacy?

Taking steps to answer these questions will provide additional valuable information on how the quality of programming funded by ACE leads to important changes in the knowledge and skills of youth that ultimately translate into academic achievement and success.

Chapter 1
Introduction and Evaluation Overview

The 21st Century Community Learning Centers (21st CCLC) program, authorized under Title IV, Part B, of the Elementary and Secondary Education Act (ESEA), as amended by the No Child Left Behind Act of 2001 (NCLB), supports the creation of community learning centers that provide academic enrichment opportunities during non-school hours for children, particularly students who attend high-poverty and/or low-performing schools.[9] The federal grants are awarded to state education agencies (SEAs), which, in turn, make competitive awards to eligible grantees to support afterschool and summer learning programs.[10] In July 2002, the federal government awarded the Texas Education Agency (TEA) $24.5 million to fund TEA's first cohort of 21st CCLC grantees for the 2003–04 school year. As of 2010–11, this and subsequent federal funding has resulted in 281 grants being awarded in Texas over seven funding cycles.[11] All centers funded by the Texas 21st CCLC program, known in Texas as the Afterschool Centers on Education (ACE),[12] 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.[13]

To ensure that grantees funded by the ACE program are positioned to achieve program objectives, TEA has developed a research-based Critical Success Model (CSM). This model includes four Critical Success Factors (CSFs), which represent behavioral changes that should be demonstrated by students and families enrolled in the program, or by the adults working on their behalf, to ensure success in meeting programmatic goals and objectives. The CSFs and corresponding milestones (key strategies that establish the foundation on which critical success factors are built) follow.[14]

---

[9] For more information see http://www2.ed.gov/policy/elsec/leg/esea02/pg55.html

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

[11] 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.

[12] 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. While 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.

[13] 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

[14] Beginning in 2009–10, with the sixth funding cycle, program guidelines were revised to require grantees to develop and implement programs in alignment with TEA's research-based CSM. Cycle 6 grantees were required to use this model to establish program goals and implement their programs. In addition, Cycle 6 grantees must collect and report performance measure data to TEA based on milestones and CSFs. Cycle 5 grantees are not subject to these requirements, although they were made

- CSF1 emphasizes both student and family engagement. Student engagement and family engagement are not necessarily achieved together, because they require different strategies and activities. The milestone for this CSF is predicated on the implementation of research-based, innovative instructional techniques and opportunities that encourage student and family engagement.

- CSF2 addresses student involvement in school, exemplified through more participation in extracurricular activities, and more mentors supporting students. The milestone for this CSF emphasizes the role of adults as advocates for students.

- CSF3 addresses the use of assessment data to evaluate and revise student activities and services. A milestone strategy is for program staff to conduct ongoing and continuous assessments to identify student needs and how to revise program services.

- CSF4 addresses staff professional development. The milestone strategy is for programs to provide all staff the "required training opportunities," which are then implemented in the afterschool program.

Table 1 presents the CSFs, including their links to outcomes, and related behaviors, performance indicators, milestones, and performance measures within the overarching Texas ACE CSM.

---

aware of the CSM when Cycle 6 was first implemented and are encouraged to use the model as a guide for improving programs.

Table 1. Texas ACE Critical Success Model

Outcomes
- Improve Academic Performance   - Improve Attendance   - Improve Behavior   - Increase Promotion Rates   - Increase Graduation Rates

Critical Success Factors (behaviors)	Critical Success Factors Performance Indicators	Milestones (grantees)	Milestone Performance Indicators
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 <u>Measurement Tool</u> - 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 <u>Measurement Tool</u> - 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 Factors (behaviors)	Critical Success Factors Performance Indicators	Milestones (grantees)	Milestone Performance Indicators
Critical Success Factor #3: Assessment Data			
• Use of assessment data to revise/reevaluate student services	• Changes in student activities following reassessment <u>Measurement Tool</u> • 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 <u>Measurement Tool</u> • 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

Many elements of the CSM can be connected to a larger framework for understanding a path to quality in afterschool programs. Figure 1 depicts a framework developed by the American Institutes for Research (AIR), which shows that youth outcomes are predicated on a complex set of interactions between several program elements, including:

- **Youth Characteristics**. Youth bring their own contribution to the afterschool setting and to their own success in that setting.

- **Community Context.** The resources and characteristics of the local and school community support the development of program goals and program design, and allow for meaningful partnerships and program guidance.

- **Program Participation**. Youth are more likely to experience benefits from afterschool program participation if they attend consistently, over time, and participate in a variety of types of activities.

- **Program Quality**. In this model, program quality is conceptualized as a series of practices and approaches that result in the creation of developmentally appropriate, high quality settings for youth at the point of service. These practices and approaches include those directly adopted by activity leaders working directly with youth, and those implemented by the organization as a whole, which craft an infrastructure to support the introduction and diffusion of effective practice in the design, delivery, and evaluation of afterschool programming.

At both the point-of-service and organizational levels, quality practices may also be categorized as supporting academic skill-building and mastery, and positive youth development. Practices supportive of academic skill-building relate to establishing linkages to the school day; having mechanisms for communicating with school-day teachers and administrators about student needs and curriculum; and having access to student data to identify student academic needs and monitor student growth. Practices related to youth development provide supports and opportunities to youth that result in a supportive, interactive, and engaging environment, as well as opportunities for youth to have choice and voice in, and ownership of, the program.

Figure 1 shows AIR's quality framework for afterschool programs. In this figure, the relationships are depicted among four components that shape youth outcomes. These are the characteristics of youth, community context, program participation, and program quality. Program quality and program participation are essential to positive youth outcomes; and, as Figure 1 shows, practices associated with quality and participation are influenced by and adapted to youth characteristics and community context. A significant goal of this report is to provide information on the relationship between the implementation of quality practices and the achievement of desired participation, academic, and behavioral outcomes.

Figure 1: AIR Quality Framework for Afterschool Programs

YOUTH CHARACTERISTICS

PROGRAM QUALITY
- ENGAGEMENT (Youth Voice & Choice)
- RELATIONSHIPS (Positive Peer & Adult Interactions)
- CLIMATE (Safe & Supportive Environment)
- ORGANIZATIONAL PROCESSES
- QUALITY AT THE POINT OF SERVICE
- OPPORTUNITIES FOR ENGAGEMENT

POSITIVE YOUTH OUTCOMES

COMMUNITY CONTEXT

PROGRAM PARTICIPATION
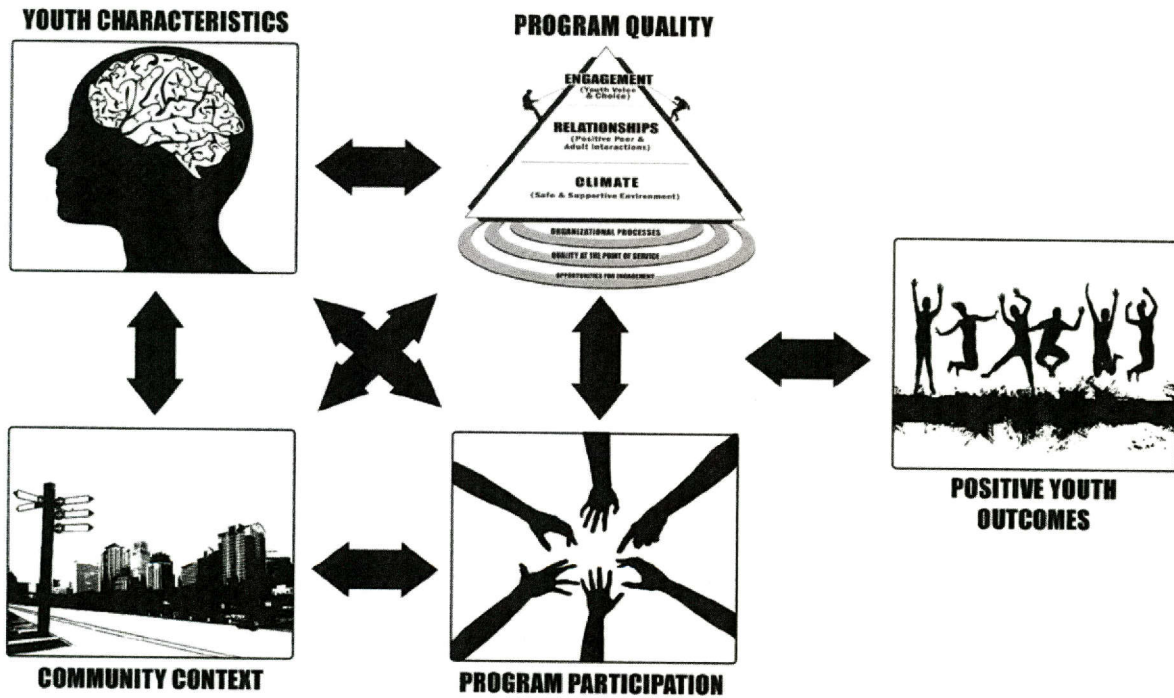
Evaluation Overview and Summary of Year 1 Findings

Through a contract with TEA, an evaluation of the Texas 21st CCLC/Texas ACE program is being conducted by the AIR and its partners, Gibson Consulting Group, Inc. (Gibson Consulting) and the David P. Weikart Center for Youth Program Quality (the Weikart Center), from January 2011 through August 2013, and possibly through August 2015 if additional funding is identified. Evaluation efforts for this contract have been largely focused on programs that were awarded grants during Cycle 5 and Cycle 6.[15] The overarching goal of the evaluation is to determine which program strategies and approaches are most effective within particular contexts in encouraging student behaviors (CSFs) that lead to improved student outcomes. As such, the evaluation is designed to address two primary research objectives:

- **Research Objective 1:** To conduct a statewide assessment of ACE programs, operations, participation, and student achievement;

---

[15] At the beginning of the 2010–11 school year, Cycle 5 programs began their third year of implementation, and Cycle 6 programs began their second year of implementation. Programs funded by the Texas 21st CCLC program are funded for five years. Applicants that received funding in previous 21st CCLC cycles are eligible to apply for funding to serve new or existing programs; however, the pre-existing grant must expire before project start date of the new grant cycle.

- **Research Objective 2:** To identify and describe innovative strategies and approaches implemented by successful Cycle 5 and Cycle 6 grant-funded programs.[16]

The Year 1 Interim Evaluation Report[17] (referred to henceforth as the Interim Report), published by TEA in April 2012, presented findings that showed the extent to which the ACE programs impacted student outcomes, as well as several critical trends in program design and implementation.

One of the findings in the Interim Report showed that student participation in ACE programs is associated with higher scores on the Texas Assessment of Knowledge and Skills (TAKS) in reading and mathematics. All students who participated in the 2009–10 program year, across all cycles in operation, were included in the impact analyses.[18] In addition, when compared with non-participating students, ACE student participants had fewer assigned disciplinary days during the school year, fewer disciplinary incidences in Grades 9–12, and fewer absences. While the findings were statistically significant,[19] the effect sizes were quite small,[20] although they were consistent with what would be expected for afterschool programs of this type (Kane, 2004).

Analyses of Cycle 5 and 6 programs operating in 2010–11 showed that across the state, the spectrum of program quality was broad, although some trends and relationships between program quality and program characteristics were identified:

- Centers staffed mostly by school-day teachers were more apt to engage in practices supportive of academic skill-building, including relying on externally-developed curricula to guide activities, developing linkages to the school day, and using student data to inform programming.

- Programs serving high school students exclusively demonstrated a lower degree of intentionality in program design and weaker linkages to school-day classes than other programs, even when staff consisted of mostly school-day teachers. However, high school program activities had higher levels of academic content

---

[16] These research objectives may be expanded to include new grantees in later years of the evaluation.
[17] http://www.tea.state.tx.us/index4.aspx?id=2908&menu_id=949
[18] While programming year in this report includes any programs offered during the summer following the school year, such extended programming has no relevance for this report. Summer programs were not observed, and student data are associated with the school year.
[19] Throughout this report, statistical significance refers to the probability that a result or relationship is random is 5% or less (p-value<0.05).
[20] Effect size refers to the magnitude of the relationship between two variables, in this case program participation and outcome. A small effect size indicates a somewhat weak relationship. Participation in the program, for students in Grades 4 through 12, had an effect size of .027 on TAKS-ELA/Reading and of .032 on TAKS-Mathematics (effect sizes are presented in standard deviation units). This is a small effect size, which indicates that on mathematics, for example, students participating in ACE programs scored higher than non-ACE students by half of one question. In general, effect sizes in educational research do not exceed 1.0.

and climate than the activities of programs serving other grade levels. This indicates that activities were focused on academic objectives, but not necessarily as a result of efforts to align sessions with specific class objectives.

- Compared to other types of program activities, academic enrichment activities had higher levels of academic content (the extent to which staff promote and students practice content-specific skills) and academic climate (a focus on developing specific academic skills, the appropriate level of challenge, facilitation of higher-order thinking skills, and feedback that promotes student learning) and more frequently embedded practices that foster youth development.

- Youth ownership—that is, youth having a role in selecting and shaping program activities—was more evident in high school programs than other programs, and student-reported engagement in program activities was higher. It is the hypothesis of the research team that student motivation plays a strong role in shaping how high school students interact with and benefit from their participation in ACE programming, and youth ownership may be a factor in motivating students to participate.

- Several characteristics increased student engagement and on-task time, and meaningful interactions between students and adults. These included: low staff-to-student ratios; structured, well-planned activities; instructors' knowledge of and responsiveness to individual student needs, interests, and personal lives; and shared social norms that guided casual interactions among staff and students. In addition, evidence suggested that the more staff adopted practices to support youth development, the higher the engagement reported by students on post-activity student surveys.

- Among elementary age students, engagement was facilitated by structured, whole-group instruction, with all students more or less focused on the same task. Among high school students, engagement was facilitated when instructors provided students with choices and responsibilities, and provided them with opportunities to use relatively sophisticated tools and materials.

Approach to Addressing the Research Objectives in Year 2

This report picks up where the Interim Report left off by describing in more detail the characteristics of program activities where student engagement was high and sustained, and the characteristics of centers themselves that support consistency in delivering high quality activities. In addition, the current report aims to make a more distinct connection between the following: center quality and student participation; center quality and student academic and behavioral outcomes; and program impact on

student outcomes, based on a replication of Year 1 impact analyses while defining student participation more rigorously.

The Year 2 evaluation was guided by seven research questions. The first four questions are associated with Research Objective 2 (which addresses innovative strategies and approaches) and the latter three questions are associated with Research Objective 1 (the statewide assessment of student results). The questions are presented in the order in which they are addressed in the report.

1. What instructional approaches are associated with high levels of student engagement at the point of service?

2. What innovative strategies and approaches can be identified from these centers that warrant replication and emulation?

3. What organizational processes are found to be drivers of instructional/point of service quality at high performing centers?

4. What is the relationship between the characteristics of individual youth, center quality, and other center characteristics and levels of student participation in ACE programming?

5. Does the impact on student outcomes vary by relevant ACE program characteristics, including center quality?

6. To what extent do students who have higher participation rates demonstrate better academic and behavioral outcomes as compared with similar students who participate in ACE programming at lower levels?

7. To what extent do students participating in services and activities funded by 21st CCLC demonstrate better achievement (along with other student outcomes) as compared to similar students not participating in the program?

Data Sources

The data sources and methods that provide the basis for the report are the interview and observation data collected on site visits in 2011 and 2012, and Texas administrative data, which includes data on ACE program characteristics, student characteristics, participation, and academic and behavioral outcomes.

2011 Site Visits

During the spring of 2011, data were collected from 40 ACE programs funded in Cycle 5 and Cycle 6 during two-day site visits. The site visit sample was selected randomly based on the following criteria:

- **Grade Level Served.** ACE programs serving elementary, middle, and high school students were represented in the sample.

- **Feeder School Performance on State Assessments in Reading and Mathematics.** Feeder schools are those attended by the students who participate in the ACE programs. For the most part, students attend programs that are provided on their school campus. Given that the focus of the 21st CCLC initiative is on improving student proficiency in reading and mathematics, selected programs spanned the achievement spectrum in these subject areas. In cases where programs served multiple schools, a weighted average was calculated to account for the proportion of ACE program participants who attended a particular feeder school during the school day.[21]

- **Rural/Non-Rural Status.** Rural afterschool programs face a different set of issues around afterschool programming than urban programs, including access to partnerships, staffing, recruitment strategies, and even program goals and objectives (Naftzger, Margolin, Kaufman, & Ali, 2006). Because of these differences, both rural and non-rural programs were adequately represented in the site visit sample.

The primary purpose of the site visits was to learn how program quality varied among the Cycle 5 and Cycle 6 ACE programs. The spring 2011 site visits allowed the evaluation team to explore program quality and the program operations and activities that differentiated *high quality* and *low quality* programs. By observing a range of programs, it was possible to identify the practices that drove quality in some programs, but were absent in others programs.

In the fall of 2011, an additional 40 Cycle 5 and 6 centers were visited, although the intent here was to visit higher quality programs identified by employing the following criteria:

- Twenty of the 40 sites were selected because they were the top ten elementary and secondary programs in terms of mean impact on reading and mathematics TAKS scores.

---

[21] The success of 21st CCLC programming is contingent upon both identifying the academic needs of participating students and crafting programming in intentional ways to specifically address these needs; and employing a service delivery approach that is developmentally appropriate and engaging for participating youth. As a consequence, the nature of programming is expected to be different from one program to another, significantly informed by the academic needs of participating students. In this regard, it was important to consider feeder school performance when constructing the site visit sample.

- Twenty of the 40 sites were selected because they had the highest average ranking on the average percentage of days attended by participating students; the average percentage of students that attended both semesters; and a balance between enrichment and non-enrichment offerings. Sites in this group also had a mean positive impact on reading and mathematics and a low to medium need for technical assistance in terms of performance on the ACE Prime Assessment, the quality assessment tool used in Texas for ACE programs.

One of the goals of the fall 2011 visits was to identify 15 ACE programs that demonstrated a capacity to deliver high quality activities. Once selected, AIR planned to visit the 15 centers a second time in spring 2012 to learn more about the strategies, approaches, and practices that supported high quality activities. As it turned out, rather than selecting the 15 centers from the fall 2011 programs, after reviewing the centers' observation data and interview data, AIR decided to select the 15 centers from the centers visited in both spring 2011 and fall 2011.

Selection of the 15 centers was based on the analysis of observation and qualitative data from the 2011 visits. Centers with high quality activities were identified based on their average observation scores.[22] A list of centers that had at least one rating in the top quintile plus no ratings below the median was created. Separately, the centers' ratings on operational dimensions (such as staff development and linkage to the school day), which had been assigned based on information from interviews and focus groups, were examined, and a list of centers with multiple high ratings was created. The centers appearing on both lists were considered, and 15 centers with multiple high ratings were identified and four alternatives were proposed as candidates for the spring 2012 visits. This list was discussed with TEA. In addition, site visitors who were familiar with the centers offered their reflections on the level of quality observed on site before AIR and TEA decided on a final list of 15 ACE programs. Table 2 presents the selection criteria for each of the rounds of site visits that were described.

---

[22] The observation protocols used were the Program Quality Assessment, for which the total score was used; the Observations of Child Engagement; the Assessment of Afterschool Practices Tool; and the academic climate scale on the Program Quality Assessment.

Table 2. Centers Participating in Site Visits and Criteria, 2011 and 2012

Round	Selection Criteria	Number of Centers
Spring 2011	Random selection, representing rural and urban, academic performance variation, and different grade levels	40
Fall 2011	20 selected because of impact on academic performance 20 selected because of higher than average student participation All sites: both academic and non-academic enrichment programming	40
Spring 2012	Prior site visit (spring or fall 2011) High observation scores Indicators of quality on organizational processes	15

The site visits conducted in the spring and fall semesters of 2011 were guided by the same basic approaches and included structured interviews with site coordinators, focus groups with staff leading afterschool activities, and observations of activities. Observed activities were rated by AIR staff using four observation protocols: the Youth/School Age Program Quality Assessment (PQA); the Observation of Child Engagement (OCE); portions of the Afterschool Practices Observation Tool (APT-O) related to supports provided by staff and tasks undertaken by students to practice specific academic skills; and the Academic Climate scale of the PQA.

A student engagement survey was also collected at the end of observed activities from participating students who were in Grade 4 or higher. Finally, surveys were collected from staff working directly with students in the site visit centers. The purpose of the surveys was to learn the extent to which staff had adopted research-supported practices to support academic skill-building and mastery, and youth development.[23]

Spring 2012 Site Visits

A key data source for the findings presented in this report were the observations and interviews conducted in spring 2012 at 15 centers identified as having the capacity to deliver high quality activities.

- Observations of three or four activities were conducted at each center visited in spring 2012, for a total of 56 observed activities. Observations used structured protocols, which were organized around specific dimensions of the activity (and are described in Chapter 2 when findings from observations are presented). Using the protocols as guides, observers created activity summaries for each session observed.

---

[23] The Interim Evaluation Report, Chapter 10, pages 104-155, describes these instruments more fully. http://www.tea.state.tx.us/index4.aspx?id=2908&menu_id=949

- Instructors of the observed sessions were interviewed following the observation. Topics included the job position of the instructor (e.g., certified teacher, youth worker—a staff person from a youth-serving organization who works directly with youth), lesson objectives, resources the instructor used for planning, and materials and resources used by students.

- Principals of the host schools, the project director from the grantee organization, and the center's site coordinator were interviewed.[24] Topics included perceived benefits for students, alignment with the school day, communication with teachers, the principal's role in supporting the program, communication with the site coordinator, and space and resources.

Data collected from the spring 2012 visits were used to develop program summaries for each center. The summaries focused on organizational processes, such as linkages to the school day, continuous improvement efforts, and staff collaboration. The summaries were analyzed to identify specific drivers of program quality. It is important to note that the 15 centers were selected first on the basis of the activity ratings from prior site visits, and second, on indicators that some of the organizational practices might support the quality of activities. Not every center was strong on all of the components that were examined to select the 15 programs, as will be evident later in this report.

Staff Surveys

In addition, surveys were completed by staff at the 40 centers visited in the spring of 2011 and these were included in the analyses. A total of 465 surveys were completed, an average of 12 surveys per center. While these surveys were the basis for much of the analysis presented in Chapter 4, to enhance the stability and power of quality estimates derived from observations and staff surveys, data obtained from the full 80 centers represented in the spring and fall 2011 samples were analyzed together. (Additional information about each of these data sources and analytic methods can be found in the report chapters where findings are presented.) Scales appearing on the survey included the following:

- Program objectives

- Creation of interactive and engaging settings for youth

- Intentionality in activity and session design

- Practices supportive of academic skill-building, including linkages to the school day and using data on student academic achievement to inform programming

- Practices supportive of positive youth development

---

[24] The grantee organization is the organization that was awarded the grant and that oversees that the grant is implemented at the centers within the organization.

- Opportunities for youth ownership

- Internal communication designed to support program development and improvement

- Training participation.

Administrative Data Maintained by TEA

In addition to new data collection activities undertaken by the evaluation, a substantial amount of information housed in TEA's administrative data systems was obtained to support the evaluation. Each of the systems, and how they were used, is described in greater detail.

TX21st CCLC Student Tracking System (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 (PPICS) database. TX21st collects data on a broad array of program characteristics, student demographics, program and activity attendance, and student outcome data (including information on student grades) directly from grantees throughout the program year. Data extracted from the tracking system were used to construct variables summarizing the staffing models employed by centers, program maturity and organization type, and levels of program attendance. Many of the variables used in analyses that assess the relationship between program and student characteristics and student outcomes were derived from TX21st.

Additional TEA Data. Both the quality-related and impact analyses described in this report used TAKS scale scores in reading and mathematics as outcomes. These analyses also included variables on student demographics, discipline incidents, school-day attendance, and grade promotion from the Public Education Information Management System (PEIMS). Campus-level performance data from the Academic Excellence Indicator System (AEIS) was used to create a non-participant comparison group.

Organization of the Report

The remaining report has five chapters, which respectively focus on quality practices, the influence of center quality, program influence and impact (two chapters), and findings and recommendations.

- **Chapter 2 – Quality Activities.** This chapter highlights the quality practices observed and documented at the 15 sites visited in the spring of 2012. Additional information is provided through a review of high quality activities observed in fall 2011. The chapter first presents features of high quality activities based on an

analysis of the selected observations conducted in 2011. This is followed by descriptions of several high quality activities observed in spring 2012.

- **Chapter 3 – Organizational Practices that Support Quality.** Chapter 3 focuses on organizational practices, first describing features of high and low quality organizational practices among the 15 centers visited in spring 2012. The chapter then presents examples of high quality organizational practices demonstrated by a sample of the 15 programs in the spring 2012 site sample.

- **Chapter 4 – The Influence of Center Quality on Student Participation and on Student Outcomes.** This chapter continues the exploration of center quality by showing its influence on student participation and student academic and behavioral outcomes among the random sample of Cycle 5 and Cycle 6 centers where site visits were conducted in spring 2011. Participation is considered in terms of intensity (how many days of participation) and duration (over how many days did participation last). Outcomes include academic achievement and behavioral outcomes.

- **Chapter 5 – Program Impact on Student Outcomes.** This chapter will be a replication of Year 1 impact analyses, but in this case, program participation is defined as occurring once a student participates for 30 or 60 days. Analyses are predicated on comparisons between participants and non-participants and between high program attendees and low attendees. Outcomes considered include TAKS reading and mathematics scores, school-day absences, disciplinary incidences, and grade level promotion.

- **Chapter 6 – Summary of Findings and Recommendations.** The findings and recommendations are intended to advise the work of the afterschool programs and TEA in delivering and supporting quality programming.

Chapter 2
Quality Activities

In the Interim Report, the quality of activities was described, but mainly in the context of the scores on observation protocols. The last chapter of the Interim Report provided descriptions of high quality activities and strategies of instructors in those activities, basing the descriptions on narratives of activities that were aligned with scores assigned during observations. Chapter 2 in the Year 2 Evaluation Report expands that discussion by describing features common to a variety of activities in which students were consistently engaged from two rounds of site visits (spring and fall 2011), and presenting examples of activities that show how the features were incorporated into the planning and delivery of the activities. These examples are from observations conducted in spring 2012.

Chapter 2 addresses Research Objective 2, which is to identify and describe innovative strategies and approaches implemented by successful Cycle 5 and Cycle 6 grant-funded programs. The chapter specifically focuses on two research questions, which were first presented in the introductory chapter of this report:

- What instructional approaches are associated with high levels of student engagement at the point of service? (Consider the CSFs reflected in the instructional approaches.) (RQ1)

- What innovative strategies and approaches can be identified from these centers that warrant replication and emulation? (RQ2)

As noted in the introductory chapter, program activities were observed in 40 centers that were randomly selected for site visits in spring 2011, and an additional 40 centers, visited in fall 2011, that were selected for site visits based on indicators that student academic and behavioral outcomes had generally improved at those ACE centers, as well as other criteria. At each center visited in 2011, activities were observed using several observation protocols that measured student engagement and other key features of the activities (e.g., interaction, supportive environment, academic focus). In the spring of 2012, 15 of the 80 centers were revisited and another round of observations and interviews was conducted.

Chapter 2 focuses first on the common instructional features of activities where students were engaged and, within the discussion, provides counter examples—features of activities during which students were not engaged. This discussion is derived from an analysis of activities observed in spring and fall 2011 and is associated with research question 1, which addresses instructional practices common to high quality activities.

The section then provides detailed information on selected activities observed in spring 2012, addressing research question 2 on strategies and approaches that might be emulated by other programs. These activities are referred to as exemplary or high quality and were selected because students were consistently and purposefully engaged. Among the seven examples presented, three are from elementary schools and four are from secondary schools. Two examples are academic enrichment activities, four are non-academic enrichment activities, and one is a homework help activity.

It is important to note that the ACE centers featured in this chapter and the following chapter were not selected because they represent all ACE centers, nor were the 15 centers observed in spring 2012 selected because they were *known* to consistently deliver high quality activities. Rather, they were observed to have done so to an extent that warranted further exploration.

Features of High Quality, Engaging Activities

Throughout this evaluation of the Texas ACE programs, differences in the extent to which students were engaged in activities have been evident. In the first two rounds of site visits (spring 2011 and fall 2011), student engagement in observed activities was measured using the PQA and the OCE. On the PQA, the centers' adjusted average scores[25] (usually over four observations) for student engagement ranged from 15.75 (low) to 64.64 (high). On the OCE, the adjusted average scores ranged from 34.21 (low) to 87.29 (high).

The features of high quality, engaging activities were examined from the spring and fall 2011 activities. Thirty-six activities were selected for the analysis. The selected activities represent academic enrichment activities, non-academic enrichment activities, and homework help activities. Selected activities include those with high PQA ratings (indicating high quality) and, for contrast, those with low PQA ratings.

One purpose of the 2011 observations was to describe what was occurring when students were highly engaged or not engaged. OCE ratings for activity segments were assigned throughout an observation. For each segment (thus, each rating), a narrative description of what the students and instructors were doing was completed. As such, each activity had an overall PQA rating for the activity, and multiple OCE ratings (accompanied by narratives) for specific segments of the activity.

---

[25] A partially-crossed method was employed when conducting PQA-related observations in the spring of Year 1 where observers were paired in an intentional manner to allow for the PQA measures to be calibrated using Many Facet Rasch Measurement (MFRM). In doing so, the evaluation team was able to obtain an estimate of whether a given rater was systematically more lenient or severe in their ratings and adjust calibrated scores to account for this systematic bias demonstrated by the rater. Additional information about this method can be found in Appendix C.

For the analysis on features of high quality activities, 36 activities were selected across three types of activities—academic enrichment activities, non-academic enrichment activities (that is, purposeful activities that build skills and knowledge and were not related to a core subject area), and homework help activities—and across different levels of quality, as shown in Table 3.

Table 3. Distribution of Observations by Grade Level and Activity Type Among Activities Observed in 2011

	Academic Enrichment	Non-Academic Enrichment	Homework Help
Elementary Centers	3 High Quality 3 Low Quality	3 High Quality 3 Low Quality	3 High Quality 3 Low Quality
Secondary Centers	3 High Quality 3 Low Quality	3 High Quality 3 Low Quality	3 High Quality 3 Low Quality

Among the high quality activities in which students were highly engaged, three features were particularly evident: clarity of purpose, intentional use of time, and an active and interactive instructor.

Clarity of Purpose

In the high quality activities, students knew what they were supposed to do and why. Instructors linked assigned tasks to the larger objective of the activity, and, going further, linked tasks and objectives to the context of the subject area and of students' lives.

One example from the 36 activities that were examined is an academic enrichment activity implemented in an elementary school. Students were to engage in a science experiment, using the framework of the scientific method to determine how much water (by volume) was contained in ice. The instructor first led a 20-minute discussion, asking students about the previous day's snowfall, eliciting students' prior knowledge related to how snowflakes are formed and what salt does to snow, and asking students to predict what salt would do if it were sprinkled on ice. This discussion set up the actual activity— to conduct an experiment—which students approached with attention and understanding. In many ways, the discussion forecasted the actual tasks related to the experiment, e.g., predicting, experimenting, measuring, and then recording results.

Clarity of purpose appeared to also be important for non-academic enrichment activities, though the purposes were not academic and were likely to be embedded in accepted professional or athletic practice (for example, the steps required to produce a film or present a play) or wide-ranging skills, such as those required for cooperative learning, effective use of technology, and effective presentation. In a fabric-dyeing class in an elementary school, for example, the instructor introduced the activity, describing

the dyes to students, explaining how they were to use the tools to complete the tasks, and letting students know how long the project would take. Similarly, in a weight-lifting class, instructors described to students the muscle groups that would be targeted during that session.

Lack of clear purpose appeared to set the stage for disruption. The examination of the 2011 low quality activities showed that activities that were not purposeful often fell apart because of students engaging in social conversation and becoming distracted. Students in both elementary and secondary schools responded poorly to seemingly purposeless activities, perhaps considering the activity to be "busywork." Often this appeared to be the result of insufficient justification for an activity on the part of the instructor. For example, in one of the activities observed in spring 2012, the topic was domestic violence, but the activity, making cards, was not related to learning about domestic violence. Students were minimally engaged in the activity. Many of them put little effort into the activity and instead talked with peers, played videos, and completed homework assignments.

Intentional Use of Time

A prominent feature of high engagement activities was the preparation and pacing that resulted in an intentional use of time. Preparation included having a planned lesson and all materials ready and distributed by the time students entered the class. An example from 2011 was a Crime Scene Investigation class. When students arrived, five tables were set with microscopes, evidence cards, rubber gloves, and other materials. The instructor formed students into five groups of three students even before the students entered the room and then directed them to stand at their appointed table.

In the high quality activities, the sequence of activities was clear (and was typically communicated to the students). While high engagement activities did not always start exactly on time, the time frame of the activity was highly delineated. This was true in both secondary and elementary center activities. Teachers made clear announcements when activities started, and when tasks needed to advance. They reminded students how much time remained and how many tasks needed to be completed in a given time frame. Students were keenly aware of the schedule and this awareness helped them organize their time and seemed to help them stay in control of the activity without causing undo pressure.

Another aspect of intentional use of time, particularly evident in the activities observed in spring 2012, was that classroom routines and behavioral norms had been established. It was evident when observing high quality activities that students were comfortable with classroom routines, because many of them followed them with only minimal reminders from the instructor. In the highly engaging activities, there was little or no time lost as students transitioned from one task to another or from one setting to

another. Behavioral disruptions, which were minimal, were addressed quickly and effectively by instructors and did not escalate or disturb the flow of the activity.

In some of the activities observed in the spring of 2012, when careful preparation and pacing were not evident, student engagement was low, and instructors had a difficult time gaining or refocusing students' attention. For example, in one class, materials were not ready at the start of the class and the instructor had to find dry erase markers and make edits to worksheets after students were in the classroom. During that time, students had begun "joking around." Minutes were lost restoring order. In a computer class at an elementary school center, the pace of the activity was described by the observer as "very slow." This was due to the instructor requiring all students to work through exercises together. Some students finished an exercise early and then had nothing to do while waiting for other students to finish.

Active and Interactive Instructors

Characteristic of the high quality activities were active instructors who continually interacted with students. This was evident in all types of activities and types of grouping (e.g., when students were in a large group, working independently in small groups, and working individually, as in homework help sessions). Successful activities were highly interactive, with the instructor continually circulating among and talking with students. In highly engaging activities, instructors did not sit at their desks. They answered questions, modeled techniques, helped with specific aspects of tasks, and questioned students for deeper understanding. They noticed successful efforts and problems, and frequently scaffolded and encouraged students. They quickly noted off-task or disruptive behavior and addressed disruptions before they escalated. In one session, even when a video was being shown, the instructor discussed what the students were viewing the whole time, commenting on the program and interacting with the students about what they were viewing.

Staff interaction was especially effective in some of the elementary school activities, because of potential social conflicts. Among the high quality elementary activities (2011 sample) were activities where students were given choices relating to the resources they wanted to use—for example, in a Board Games class, elementary students chose among games and a craft project. In this setting, instructors circulated and interacted with students in part to maintain fairness and to help students decide how to resolve any conflicts that might arise.

An interactive instructor appeared to be particularly essential for homework help sessions, during which students often worked alone or in small groups. In the high rated activities at both the elementary and secondary levels, homework help was a highly interactive experience, in which the instructor (or instructors) and students were constantly engaged with each other, discussing problems, plans, prior knowledge, and other content-related ideas.

By contrast, in the low scoring sessions, homework help was set up as an opportunity for individual work, and there were only minimal efforts on the part of instructors to actively engage with students. The result was that many more students were off task, with students appearing to view homework help as a break from the school day, unless staff actively guided them and encouraged them to stay on task. In one session where student engagement was low, the instructor was active, but because there were so many students (16), he or she could not interact with all of them, and spent time with only a few. The result was that some students were off task and the general environment was described as "very unproductive." In a high school activity, students were expected to conduct independent online research on colleges. The instructor arrived late and repeatedly left the classroom during the session. The majority of students were off task, somewhat randomly surfing the Internet and watching YouTube videos.

Other Factors

Other characteristics of high quality activities where students were consistently engaged were closely aligned to the three factors mentioned above: clarity of purpose, intentional use of time, and an active and interactive instructor.

- **Student Choice.** Successful activities in the 2011 sample, even those that were highly prescribed, provided a fair amount of student control of the activity through specific choices. Low engagement activities demonstrated a noticeable lack of choice. Limited choice was often followed by disruption as students seemed to struggle to find a reason why choice might be limited in an activity that was not driven by clearly understood academic or other purposes.

- **Separating snack time from class time.** Snack time disrupted activities when not separated from work time. The activities observed in 2011 indicated that having no boundaries between snack time and work time was confusing for students. In some of the activities where students were most engaged, snack time preceded the class time, with class delayed by a few minutes (as many as 17 minutes in one case), but students started work promptly when snack time ended. In a homework help session, students were asked to do their homework while snacks were distributed and eaten. It took students more than a half an hour to become engaged in the homework activities.

- **Variety.** Successful elementary activities (2011) contained a great deal of variety throughout the session related to the overarching objectives of the activity. This was also evident in the activities for young students that were observed in 2012.

Summary

This section addressed the first research question: What instructional approaches are associated with high levels of student engagement? In examining the activities during which students were consistently engaged, the instructional approaches observed in different types of engaging activities were to direct instruction toward a clear purpose, to use time well, and to be active as well as interactive through the activity. Two of these approaches (clarity of purpose and intentional use of time) show the importance of planning. The last approach suggests the importance of the instructor relating to the students and providing oversight and guidance. A summary of specific activities and practices associated with high and low engagement by activity type and grade level can be found in Appendix A.

The next section presents examples of activities where students were consistently engaged. The examples illustrate different strategies instructors employed to embed purpose, planning, and an active and interactive instructor in different types of activities.

Examples of High Quality Activities

The second research question is: "What innovative strategies and approaches can be identified from the 15 centers (where site visits were conducted in the spring of 2012) that warrant replication and emulation?" To answer this question, examples of high quality activities of different types are presented in this section. In these activities, the instructional approaches described earlier are evident. Clarity of purpose, for example, is illustrated by the focused learning objectives that anchor both academic and non-academic activities. Intentional use of time is evident in lesson plans, established routines, prepared materials, articulated next steps, and reminders about how much time remains to complete tasks. All of the instructors in these examples were active, meaning that they taught and guided activities; and all were interactive, meaning that they related to students throughout the activity.

As noted earlier, the 15 centers where activities were observed in spring 2012 were selected from the centers where site visits had already been conducted. As Table 4 shows, 56 activities were observed at the 15 centers. Nineteen were academic enrichment sessions (12 of which were observed in elementary centers), six were homework help or tutoring sessions, 27 were non-academic enrichment sessions, and three were sports/recreation sessions. Observed activities were fairly equally distributed across elementary schools, where 27 activities were observed in seven centers, and secondary schools, where 29 activities were observed—seven in two middle schools, and 22 in six high schools.

The types of observed activities differed by grade level, as shown in Table 4. In high schools, only three academic enrichment activities were observed, two of which were in

one center. As interview respondents indicated, some high school centers did not offer academic enrichment sessions and some dropped such sessions for a semester at a time. Only in high schools were sports and physical education sessions observed. In addition, in high schools, activities focusing on youth leadership (among the non-academic enrichment activities) were offered in several schools.

Table 4. Types of Activities Observed by ACE Programs Serving Different Grade Levels in Spring 2012

	Academic Enrichment	Homework Help/ Tutoring	Non-Academic Enrichment	Recreation	TOTAL
Elementary	12	2	13	0	27
Middle	4	0	3	0	7
High	3	4	12	3	22
TOTAL	19	6	28	3	56

Source: Activity Narratives Spring 2012

Apart from the type of activities observed, two other characteristics distinguished activities by grade level setting. The majority of the high school activities were instructed by a certified teacher (of 22 activities for which the type of instructor was recorded, 14 were instructed by a certified teacher). Thus, in this sample, even though there were few academic enrichment sessions provided in high schools, most sessions were facilitated by a certified teacher. In contrast, more elementary activities in the sample were instructed by youth workers (14 of 26 activities for which information was recorded) than by certified teachers (eight activities).

High school and middle school activities had fewer participants than activities observed in elementary schools. In the high school centers, of 22 activities observed, six or fewer students participated in nine of the activities. In the middle school centers, six or fewer students participated in three of seven observed activities. However, in elementary school centers, six or fewer students participated in only two of the 27 observed activities.

Observations were conducted using protocols that identified several key dimensions of the afterschool activities. Foremost was student engagement, for which observers noted whether all or nearly all students were engaged in the activity and whether engagement was sustained throughout the session. Another dimension was content, and the observer learned, both from the observation and a post-observation interview with the instructor, what the learning objectives were, the topics and skills covered, materials used for the activity, and resources the instructor used to plan the lesson. Observers also paid particular attention to the organization and pace of the lesson and to the

instructor's effectiveness in keeping students active and on task. Student activities were described—including what students were doing, and how much they were interacting with other students and with materials. Another dimension covered was opportunities for students to exercise choice and apply skills such as leadership skills, cooperation, communication, and problem-solving.

Of the 56 observed activities at the 15 centers, 24 were assessed as high quality activities that successfully engaged students in purposeful learning. For each of these activities, observation narratives show that each of the classroom dimensions described above was met at a high level. Among the 24 activities:

- Nine sessions were academic enrichment sessions. All were observed in four of the seven elementary school centers, none in secondary schools.

- Fourteen sessions were non-academic enrichment sessions. Six were observed in elementary school centers, one in a middle school center, and seven in high school centers. The sessions were offered in nine of the 15 centers.

- One session was a homework help session observed in a high school.

Seven examples of high quality activities, during which students were consistently engaged, are presented in this section. Among the examples are two academic enrichment sessions, both offered in elementary centers; these sessions are highlighted because they were engaging and purposeful, and because they were structured differently from one another. Four examples of non-academic enrichment activities are presented; they were observed in both secondary and elementary centers and represent a range of topics. Finally, one homework help session is presented. This was observed in a high school.

Examples of High Quality Activities – Academic Enrichment

Academic enrichment activities are presented first because they are essential features of the 21st CCLC programs and they directly address 21st CCLC goals related to improved academic performance. For each activity, key features are highlighted. These features include specific examples of instructional approaches that contributed to student engagement and learning, and that may be emulated in other ACE programs.

Example 1 – Academic Enrichment Session in Adams Hill Elementary School (Northside Independent School District). The first activity that is featured is a language arts activity that was observed at Adams Hill Elementary School. Eighteen fourth-grade students participated. The center is operated by the Learning Tree Program, which distributes curricula developed by Northside Independent School District (NISD). The activity was based on a "high level" lesson plan developed by the NISD for the Learning Tree Program. The lesson plan was given to the instructor two weeks before scheduled implementation so that the instructor had time to adapt the

lesson to meet the needs of the students. The lesson plan clearly articulated the learning objectives of the activity, which were to reinforce student knowledge of plural nouns and verbs, and improve their sentence structure skills. The lesson objectives were aligned with the regular school-day curriculum, and student activities were adapted for the afterschool environment.

The instructor began by clearly stating the purpose of the first activity: "We are going to talk about plural nouns and verbs." To learn about students' prior knowledge, the instructor told students they needed to "get a true definition of nouns and verbs," and asked students to provide the definitions. Following this, the lesson consisted of three parts, each of which used a different grouping pattern and engaged students in different activities: answering questions in a whole group setting, working in pairs during a peer review segment, competing with one another in a team format, and working individually on journals.

The session was extremely well organized, and little time was lost, even as students transitioned from one group setting to another. The lesson started at the scheduled time, and all materials were organized and available. Routines had already been established, as evidenced by students knowing how to move between settings and transition to new activities. The sequence of activities was logical—for example, beginning with discreet skills and ending with a more holistic writing activity. Pacing was such that all students were busy for each of the lesson segments. The instructor helped students pace their work by calling out when three minutes, two minutes, and one minute remained to complete an activity. In the few occasions when a student was off task, the instructor was effective in redirecting behavior in a positive manner. At no point did behavioral problems interfere with the pace of the activity.

The activity provided several opportunities for students to interact with one another. Students engaged in a peer review of their work and participated in a game. In one activity, communication skills were called upon when students responded to questions and read sentences aloud. During the journal exercise, which concluded the activity, students applied new knowledge and skills to their written entries, and engaged in a form of self-reflection, writing about themselves or their day.

The instructor provided high quality feedback throughout the lesson. She did so by circulating among students as they worked and checking in with them to ensure they understood the assignments. When students had questions, the instructor used effective scaffolding techniques designed to gradually remove support until the student felt autonomous, which contributed to students' answering their own questions and solving difficulties. The instructor also continually encouraged students with remarks such as, "I'm impressed" and "That's pretty good."

At the close of the session, the instructor checked student learning, telling students who sat in a circle as the activity drew to a close, "I am going to time you on how to do

everything I just taught you." She asked students to say a word (e.g., a single noun) and then chose someone to give an answer (e.g., the plural form of the noun). This was done in an informal setting.

Example 2 – Academic Enrichment Session in Garriga Elementary School (Port Isabel Independent School District). The second activity is also a language arts activity that was observed in the Garriga Elementary School center. Ten first-grade students participated in the activity. The activity was taught by the school's literacy intervention specialist, who brought specialized expertise and resources to the session.

The purpose of the lesson was to teach the students how to read, spell, define, recognize, and use eight new vocabulary words. The instructor used numerous resources to prepare for and deliver the lesson, including Fry's Word List of high frequency words, the PRIME blueprint for ACE lesson planning, a summary of a story students were learning in their school-day classes, and narrative texts from which the instructor selected readings. During the activity, instructional resources included flashcards, strips of paper with the words placed in small plastic bags for each student, an LCD projector, and Fry's word lists.

The instructor introduced the activity and told students the purpose of the lesson. Fitting for a vocabulary lesson, the instructor used complex words in her introduction, saying, "I want you to be able to read these with automaticity and accuracy—that means right away and correctly." Students were told that knowing these words was important for understanding day-to-day conversations and texts they will read—thus, relating the lesson to students' lives. After this introduction, the instructor and students quickly went through the list of words, with students repeating the words as the teacher held up flashcards.

In this session, the activities and settings were mixed so that time spent in whole group activities was interspersed with brief exercises. For example, after a few minutes in a whole group setting, the instructor asked each student to work with a partner during short, timed exercises, some lasting as little as one minute. In one short exercise, students, working in pairs, were asked to "whisper to your shoulder partner what animal group a bear belongs to. Don't say it out loud – whisper! You learned this in science, I know you did." When a student responded correctly, the instructor said, "Did y'all hear what Albert said? Mammal! If you agree, put your thumbs up." Another exercise was called "showdown," during which students were asked to find the flashcard in their packet with the written word that matched the word the teacher had said. "I'm going to say the word and I want you to show it to me. If you're holding this word, then you're right." In another exercise, students used a projector to complete a sentence projected onto a screen.

Throughout the session, the instructor gave students feedback and scaffolded student learning, although because of the format and pace of the activity, there was little time to engage in extended feedback loops.

Even though the session was very much directed by the teacher, the students were allowed to move around as they wished. For example, some of the students chose to sit in an inflatable whale pool, and others on a bench. Typical of structured activities, there were few other options for student choice.

At the conclusion of the observed session, the instructor told the students the purpose of the next session, planned for later in the week, and discussed the theme of one of the stories, briefly describing the plot and asking the students questions about how that plot might relate to their own lives.

Examples of High Quality Non-Academic Enrichment Activities

Non-academic enrichment activities offer a great deal of diversity to afterschool programming and are authorized by the federal 21st CCLC legislation. The topics and disciplines covered include arts-related activities (themselves reflecting a high degree of variety), and activities that address such diverse knowledge and skill areas as culinary arts, leadership, nutrition, carpentry, and jewelry-making. The activities frequently are offered because of the interest of students, whom many of the ACE programs survey or informally ask about activities in which they are interested.

The following examples of non-academic enrichment activities had at least six participants. Those with fewer participants (one activity had only two participants for part of the session) were excluded because the instructional approaches used may not be applicable to activities with more participants.

Example 3 – Film Class at Smith Elementary School (Del Valle Independent School District). The film class at Smith Elementary School was attended by 10 fourth- and fifth-grade students. In this observed activity, students continued their work on producing a film. A key feature of the project was that the students were responsible for all creative and logistical components. The movie was both written and produced by the students. The instructor told them they could not create a film about "zombies or vampires," but other than that, the students were given free rein on film content. Students wrote the plot and the dialogue, directed and acted in the film, developed the production schedule, managed the props and costumes, and designed and hung film posters around the school to announce the film's release. During the observed activity, the students applied skills learned in previous sessions while continuing their work on the film project.

Engagement was high throughout the activity, with some fluctuation among students who only briefly did not have an active role. Even though the students had been working on the film project since February, two months before the site visit, it was evident that

they had not lost interest in the film and they spoke excitedly about the upcoming wrap party and film premier.

Several tasks were planned for the session: filming a serum-drinking classroom scene, filming additional footage to insert into a fight scene, and choreographing another fight sequence using the 'green screen' (importing a background scene on which a film sequence is overlaid). The session (and likely the preceding sessions) provided students numerous opportunities to practice high-level skills, such as cooperation (students worked together on all aspects of the film), creativity (students designed choreography and scenes), technology (working the camera and figuring out how to use the green screen to film a fight), and problem-solving (how to film a fight in a classroom).

The observed activity was fast paced, and students accomplished the tasks the instructor had planned. It was evident in this activity that the students had already developed established routines for setting up and shooting scenes, using time efficiently to accomplish their tasks.

The instructor expressed consistently high expectations and treated students with professionalism. The instructor worked constructively with students who were flubbing lines, asking them questions such as: "What part is confusing to you? Is it 'tell me?' Would you rather say 'say'?" He frequently praised students for their work—for example, on their improvising, catching problems related to the continuity of the scenes, and taking direction. Nearly always, the instructor offered recommendations that were backed by explanations and illustrations—for example, asking extras to be quiet for a moment while filming dialogue, or telling the student director how to crouch down and then rise to film a rising shot with a smooth motion. He was patient when students did numerous takes of a scene in order to produce the right footage. Students also showed patience with the filming process, doing many retakes in an effort to get the scene right. It appeared that patience and attention to detail were exemplified throughout the project. The atmosphere of the session was positive, and the instructor and students interacted with one another in a relaxed and respectful way.

Example 4 – Media Arts Class at Dunbar High School (Fort Worth Independent School District). The Media Arts session observed at Dunbar High School and highlighted in this section was attended by nine high school students. The instructor was a community member who works in the sound industry. The observed activity was one in a series on the creation of sound in movies, and focused on how sound is transmitted, and how computer programs can be used to create and alter sound. The activity had three main parts. It began with a review of the theory of sound, during which the instructor explained how acoustic energy transmits to electrical energy. In the second part of the session, the instructor showed a video from "Charlie's Angels," and demonstrated that all of the sound in the video was created in the studio. In the third

part of the session, students applied what they learned; a student first sang into the microphone and then the students used the computer to practice layering sound and altering the student's voice.

The students were amazed by the video clip and the live demonstration. The teacher encouraged students to think deeply and apply their knowledge, asking them questions that required prediction and analytic thinking. Questions included, "What do we do first?" "Could someone describe what compression is for?" "What can we do to enhance her voice?" "And you would hear more of this kind of change if we did what?" The instructor continually guided students as they built and integrated knowledge and skills and considered applications to real world situations. Such practice enhanced the purpose and the relevance of the lesson.

As in the other sessions described in this chapter, the activity was well paced and time was used productively. All materials were ready, and technology equipment was set up and functioning by the time the class began. Students appeared to know and adhere to class routines, and there were no behavior incidents that disrupted the flow of the activity. The sequence of the session (lecture, demonstration, experimentation) worked and no time was lost moving from one type of activity to another. Throughout the session, the environment was congenial, and students and the instructor demonstrated warm and respectful behavior toward one another.

Example 5 – Theater Project at Maria Moreno Elementary School (Big Thought/Dallas Independent School District). The theater class at Maria Moreno Elementary School's center was attended by 13 fourth-grade students. The instructor, who was assisted by a college student, was a community member who directed plays in the community. The activity focused on students creating and memorizing dance choreography. The dance included a rap routine that was based on mathematics concepts students learned during their school-day class.

The class began with a warm-up routine, in which everyone stood in a circle and sang "I feel so good." The teacher then pointed to a student, who called out a body part (e.g., legs), and then, on cue, all of the students shook their legs to the beat as they sang, "I feel so good in my legs!" Students then engaged in stretching exercises. Following the warm-up, students moved to the stage and practiced the dance they had created in the last session, following the eight-count pattern that is typically used to organize choreography and learn routines. The instructor then asked students to come to the front of the class, one at a time, to teach the other students a new dance move that would be added to the routine they just practiced. After this, students worked in small groups to practice the dance. Each group had a dance captain, selected during the prior session, who knew the routine students had learned in earlier sessions and helped other students learn the new routines. The dance captain modeled dance moves and helped students who did not know the dance well. To conclude the activity, the class

performed the dance together, and then broke again into small groups, each group performing the dance for the rest of the class. The dance captain for the next session was selected at the end of the session.

In this activity, it was evident that students had learned a sequence of routines related to dance and choreography. For example, they sang and warmed up with stretching exercises prior to engaging in choreography and dancing. The group routines (e.g., election of dance captains, role of dance captains, adding new choreographed steps) were understood and followed. The instructor was an active facilitator, sometimes asking students to explain their thinking, and allowing students ample opportunity to create and demonstrate what they had learned. Frequently, the instructor stood near students as they worked in small groups. At the end of the session, the instructor told each student what he or she did particularly well during the activity and then gave each student a congratulatory sticker. Students smiled and laughed during the warm-up exercise and continued to smile while practicing the dance. They clapped enthusiastically for one another.

Example 6 – Leadership Class at Rivera High School (Brownsville Independent School District). A leadership session offered at Rivera High School was described as "an exceptional session" by the observer. Eighteen high school students attended the session, which was taught by a certified teacher. The session was based on the United States Air Force curriculum and lesson plan, supplemented by information from the Internet, and informed by the instructor's own experience. The instructor prepared a PowerPoint presentation to organize information and bring up key points. The focus of the entire session was working cooperatively in teams.

The instructor was very familiar with the material, and paced the session appropriately. All the materials, including a leadership book for each student, were distributed by the time the session began. There was no down time in the session, which began and ended with students performing a military salute and a chant.

Unlike many of the activities featured in this section, this session was entirely delivered in a whole group setting. Within this setting, students were engaged in a sequence of learning new information, responding to questions, and engaging in lively discussion. Many examples provided by the instructor related to students' lives, and students provided relevant information about their lives throughout the discussion.

Students were alert and engaged in the conversation, volunteering to answer questions and providing relevant personal information throughout the session. When students did not know the answer to a question posed by the instructor, or when they were slow to respond to questioning, the instructor encouraged students to respond with statements such as, "you're on the right track," and also had other students assist each other in responding to questions. Most students actively volunteered to respond to questions and contributed to the discussion.

The entire session was very positive. The instructor and students were energetic and had matched affect during the entire session, and students were respectful of one another and the instructor. The instructor responded encouragingly to students throughout the discussion. Students actively listened to each other and the instructor for the entire session.

Homework Help

Few homework help sessions were observed. This was due in part to time constraints around the site visits, which allowed only four activities to be observed in the course of a two-day site visit, and in part to more attention being directed toward academic enrichment and non-academic enrichment sessions. The example presented here was observed in a high school.

Example 7 – Homework Help at Connally High School (Pflugerville Independent School District). At Connally High School, a science tutorial attended by three students was observed. The instructor was a school-day science teacher. There was no specific content for the session and it was open to all grades. Students came to get help with science assignments and to complete any lab work that was unfinished. Students were not required to stay for any set period of time and were allowed to move on when they had completed an assignment or activity. The session lasted one and a half hours.

All materials were ready when the students entered the classroom, and quiet classical music was playing. Students entered the room quietly and immediately began working. They had gotten their own assignments, and were responsible for completing their work (the day's science assignment). As the students worked, the science teacher frequently asked students if they needed help, provided materials, and explained problems when asked. There were several brief exchanges between the instructor and students related to the assignment, with the teacher moving near the students he assisted.
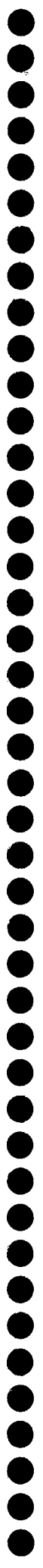
Throughout the session, students were focused on their work and there was no laughing or talking about other topics. The session climate was focused and very respectful, with the teacher greeting and saying goodbye to each student, and students using respectful language.

Summary

A major purpose of the evaluation was to identify the strategies and practices used during activities where students were consistently engaged (research question 1), and to then describe the strategies observed during high quality activities that could be replicated in other sessions (research question 2). First, the features of high quality activities were identified from observations conducted in 2011. Examples were then provided from observations conducted in 2012. Each of the examples show how the features of high quality activities—clarity of purpose, intentional use of time, and an

active and interactive instructor—were key features of the activities. In numerous sessions, the purpose was highlighted and students knew exactly what they were expected to do and how it corresponded to the sessions' purpose. The sessions were well paced and preparation was evident from the moment students entered the classroom, with materials distributed and instructors ready to begin the session. Students were comfortable with classroom routines and this resulted in no or minimal downtime when they moved from one activity to another. The teachers were highly interactive with students, moving among them, attentive to their needs, and responsive. Each of these features resulted in a high level of engagement and also a high level of productivity observed during the session. Essentially, goals were accomplished and/or projects moved forward. For example, dances were developed, focused and clarifying discussions were held, decisions were made, and homework was completed.

The next section focuses on the third research question, "What organizational processes are found to be drivers of instructional /point of service quality at high performing centers?" The section examines several organizational processes and characteristics that distinguish the centers where site visits were conducted in spring of 2012.

Chapter 3
Organizational Processes

The previous chapter examined program quality at the activity level by focusing on activities where students were engaged throughout, and comparing these activities to those where students were not as engaged. The chapter pointed out that clarity of purpose, intentional use of time, and an active and interactive instructor were characteristic of the activities where students were engaged; and that when these characteristics were absent, even for a time, students were less focused, more passive, and more likely to engage in off-task behaviors.

One of the challenges in analyzing afterschool programs is teasing out whether the high quality activities are driven by organizational processes, such as staff development and monitoring for improvement, or whether individual instructors drive quality. That is, what drives quality: is it a gifted and prepared instructor or a group of instructors who are guided and supported by program leaders, guidelines, and processes? An assumption is that if the instructor is acting alone (that is, without consistent program support), activities will not be consistent in their quality. One instructor may plan and deliver a well-conceived and well-delivered series of activities, while another may not.

Quality was examined at the organizational level in order to address the third research question: "What organizational processes are found to be drivers of instructional /point of service quality at high performing centers?" At each of the centers visited in spring 2012, the site coordinator, school principal, and the project director from the grantee organization were interviewed. Brief interviews were conducted with the instructors of the observed activities to learn more about the activity, and about organizational supports related to developing and delivering activities.

The interview data were the basis for the program summaries, which were created for each of the 15 centers. The summaries were based on 12 program dimensions, grouped in three categories, depicted in Table 5. The dimensions were identified over the course of the three site visits to ACE programs (2011 and 2012) because they were observed to be important components of a structured and purposeful ACE program, and were consistent with the 21st CCLC program model, as well as the Texas CSM. Categories and the related dimensions are presented in Table 5.

Table 5. Categories and Dimensions on 2012 Program Summaries

Category	Category Description	Dimensions Associated with the Category
Center Intentionality	Purposefulness in defining and supporting center activities	• Academic Enrichment Activities • Homework Help Activities • Non-Academic Enrichment Activities • Student-Centered Programming
Organizational Practices	Structures and processes that support high quality program delivery	• Staffing • Links to the School Day • Monitoring for Improvement • Staff Collaboration and Staff Development
Community Connections	Connections the center has made to the community, as evidenced by contributions the community has made to the center	• Community Connections

- **Center Intentionality.** Analyses conducted in Year 1 of the evaluation examined activity-level intentionality. In Year 1, at the activity level, intentionality referred to embedding academic content in afterschool activities in a deliberate way. The analysis of the qualitative data collected in the spring of 2012 goes further, in that center-level intentionality, which refers to a center's purposefulness in defining and supporting its activities, was also examined. In particular, center intentionality was examined to determine whether or not the center had clearly articulated a model for its various types of program activities—specifically for academic enrichment activities, homework help activities, and non-academic enrichment activities. In addition, center intentionality related to student-centered programming (programming that supports youth development, gives students both choice and voice, and/or provides youth with leadership opportunities across program activities) was assessed.

- **Organizational Practices.** Organizational practices are a set of consistent practices that are in place to ensure activities are high quality and engage students. Among the spring 2011 and fall 2011 centers that were visited, there was high variation on several dimensions of organizational practices—particularly staff development, staff collaboration, and monitoring for improvement. There was less variation on linkages to the school day.

- **Community Connections.** Community connections was another program category recognized as important in the 2011 site visits. Community connections

established the potential for programs to bring in community resources (such as experts in certain areas, and material resources) to their programming. This was one of the dimensions that was considered a potential hallmark of high quality centers.

Each dimension for each program was rated as high, moderate, or low in terms of quality, based on a rubric specific to each dimension. Ratings were assigned by the site visitors, and then checked by another staff member to determine if the evidence supported the rating.

It is important to note that the centers that were included in the spring 2012 visits were selected based primarily on the quality of *activities* observed on the spring 2011 and fall 2011 site visits, and, to a lesser extent, because multiple operational dimensions were rated high, based on interviews conducted while on the 2011 site visits. The number of high ratings varied from as few as two to more than 10 per center. (There were more high ratings among the centers that were visited in fall 2011 than spring 2011, in part because of the site selection, and in part because the information gathered on site was more targeted and provided more evidence for a high rating.) Also noteworthy is that the sample size is too small, and the number of observed activities too few per center, to conclusively state that the organizational processes drive or do not drive point of service quality. There are, however, numerous examples of high quality organizational processes that will be featured in this section, and, when possible, linked to the quality of activities.

Center Intentionality

Center intentionality refers to the center's purpose, and the degree to which the purpose is clearly defined and appropriate resources directed toward fulfilling that purpose. Just as activity-level intentionality is a purposeful focus of time and resources on achieving targeted learning objectives within the activity, center intentionality is a purposeful focus of time and resources on the different types of activities the center provides.

Intentionality was examined across three programmatic areas: homework help, academic enrichment, and non-academic enrichment. As Table 6 shows, there were differences in the degree of intentionality by grade level and by the type of activity.

Table 6. Center Intentionality Among 15 ACE Centers – Spring 2012

Center	Homework Help	Academic Enrichment	Non-academic Enrichment
Elementary School Centers			
A	High	High	High
B	High	High	Moderate
C	High	High	Moderate
D	High	High	Moderate
E	High	High	High
F	High	High	High
G	High	High	High
Middle School Centers			
H	Moderate	High	Moderate
I	High	Moderate	Moderate
High School Centers			
J	Moderate	Moderate	High
K	High	Moderate	Moderate
L	High	Low	Moderate
M	Moderate	Low	Low
N	High	Moderate	Moderate
O	Low	Low	Low

Source: Spring 2012 Program Summaries

- All of the elementary school centers in the spring 2012 evaluation study were rated high on intentionality related to academic enrichment. One middle school and none of the high school centers were rated high on intentionality related to academic enrichment.

- All of the elementary school centers in the spring 2012 evaluation study were rated high on center intentionality related to homework help. Half of the eight secondary centers, including one middle school center and three high school centers, were rated high on center intentionality related to homework help.

- Center intentionality related to non-academic enrichment was mainly exhibited by elementary centers visited in spring 2012, with four of the seven rated high. Only one secondary center, a high school, was rated high on center intentionality related to non-academic enrichment activities.

- Among the eight observed secondary centers, none had more than one high rating on intentionality, and two had no high ratings.

Each of the centers showing high intentionality for academic enrichment committed resources to developing and/or obtaining structured lesson plans. Five of the eight centers developed their lesson plans for the academic enrichment activities, usually with staff members working in teams to develop and/or review the plans. Two centers, including the only secondary center in this group, used lesson plans that were developed at the district level, and one of the centers mainly purchased curricula and lesson plans from different vendors. (Note, however, that the latter center defined non-academic enrichment activities as academic enrichment activities—an issue that will be discussed shortly—and saw distinctly similar intentions between the two types of activities. The center's perspective is evident in its choice of vendors, which included Campfire USA, the Theater Action Project, and Green It.)

For homework help, approaches in the elementary centers were similar across centers. With one exception, all elementary centers required all ACE program participants to attend homework help sessions. The only exception was an elementary center where only students identified as struggling by the Title I coordinator or a teacher were required to attend the homework help sessions. Typically, in the elementary centers, students worked in a large group in the cafeteria to complete their homework as instructors mingled and supported students as needed. Site coordinators at most of the elementary centers described a "no worksheet" policy. When students had completed their homework, they had other activities available, such as mentoring their peers, reading, or playing games. The centers each had ways for instructors to learn about student academic needs and homework expectations. For example, in several programs, each student had a folder containing daily homework assignments. In one elementary center, the site coordinator worked with the academic coach, who gathered homework assignments and lesson plans from the school-day teachers, and made copies for the instructors who facilitated the homework help sessions.

The elementary and the middle school centers where intentionality related to academic enrichment was high demonstrated similar perspectives on what an academic enrichment activity was. The activities were curriculum-based, for example, and lesson plans addressed specific learning objectives.

Among the centers demonstrating lower intentionality on academic enrichment programming (one middle school center and six high school centers) and homework help, several definitions of these types of activities were offered by site coordinators. When asked to describe the centers' academic enrichment activities, four of the high school site coordinators referred instead to tutoring sessions, including one site director who also mentioned the school's credit recovery courses. Two high school site

coordinators referred to non-academic enrichment activities rather than academic enrichment activities.

When asked about homework help activities, two high school site coordinators referred to tutorials and said that the tutorial sessions in the schools, which are extensions of the school-day classes and facilitated by classroom teachers, are provided outside of the context of the ACE program; the ACE program mainly provides supportive infrastructure, particularly snacks and transportation.

Among the 15 ACE centers, intentionality related to non-academic programming was mixed, with the majority of centers (eight) indicating a moderate degree of intentionality, two indicating low intentionality, and five indicating high intentionality. One reason for such a mix of ratings is that non-academic enrichment programming appears to be understood differently among the centers. In fact, given the responses of site coordinators, the term non-academic programming (rather than non-academic *enrichment* programming) more accurately reflects the understanding of many of the interview respondents.

In several centers, site coordinators, and in some cases project directors, referred to non-academic sessions as opportunities for fun, physical exercise, and socializing, rather than enrichment. Of three high school coordinators, one described the sessions as "fun enrichments" where "students are allowed to be themselves"; another as providing time for students to break from their rigorous academic demands and participate in sports; and another as "time for students to unwind." This viewpoint was expressed in two elementary school centers as well, with one site coordinator adding that the non-academic programming (which was recreational) was a way to "hook the kids in the program" so that they would attend the academic sessions. In another center, the site coordinator did not distinguish between non-academic and academic enrichment activities, because the non-academic sessions incorporated some content and skills related to school-day courses.

One of the site coordinators, who placed a high importance on non-academic enrichment activities, described them as the "most important" part of the afterschool program. The site coordinator said non-academic enrichment activities expose students to learning associated with the arts or other professions. The site coordinator said, "If kids don't have experiences, they don't have anything to aspire to because they don't know. Many of our kids are at risk. They are economically disadvantaged. They have language issues for whatever reason and so our business and our focus is to provide hands-on, intentionally programmed [activities]."

Examples of Center Intentionality – High and Low

In response to the research question, "What organizational processes are found to be drivers of instructional/point of service quality at high performing centers?" several examples of center intentionality are presented. These include two examples of high intentionality, one related to academic enrichment and the other to non-academic enrichment, and one example of low intentionality.

Example 1 – High Intentionality on Academic Enrichment. As noted earlier, all of the elementary schools and one middle school where the 2012 site visits were conducted focused on academic enrichment. The respondents at the Adams Hill Elementary School described in detail a very focused approach to academic enrichment programming. The center was associated with the Learning Tree Program, which develops afterschool academic enrichment curricula for the ACE programs under its purview. The curriculum was developed by teachers and reviewed by a curriculum committee, which included site coordinators for the ACE programs. The curriculum and lessons plans were modified as campus and program needs evolved. The Learning Tree Program project director said, "We conduct a needs assessment in June. That reveals any needs of the campus, even though the campuses we serve focus on the basics—reading, math, science, and social studies."

The curriculum followed the scope and sequence of core classes in the district, although, as the site coordinator said, "Our curriculum is developed to be more hands-on. There are activities. There are games. This is because we all understand that students have been in the classroom for so long that they're going to be bored. The curriculum is definitely geared towards a hands-on approach, or going away from just sitting down in a lecture style. That's what we work to stay away from."

Example 2 – High Intentionality on Non-academic Enrichment. At Connally High School, two high quality non-academic enrichment activities were observed. The non-academic enrichment activities—described as the "main feature" of the ACE program by the site coordinator—followed a "club" format, in which students were engaged in the learning activities and assumed a share of responsibility related to planning and, in some cases, securing future funding for the club. All activities were expected to have students work toward culminating events or projects. This was evident in the activities that were observed.

The selection of activities was based on a campus needs assessment. Although high school teachers proposed a specific activity (or club) to lead, the site coordinator made the final selection. The site coordinator said, "We look at the campus plan and goals and see how we [the ACE program] can tie enrichment and academic activities into that." The activities themselves were delivered by high school teachers, who, the site coordinator said, "connect well with kids and already have very good ideas about programming and things they want to do," and by skilled community members who were hired to teach specialized activities, such as digital media and mariachi band.

Example 3 – Low Intentionality on Non-Academic Enrichment. In another secondary center, two high quality non-academic enrichment activities were observed, but center intentionality related to non-academic enrichment activities was assessed as low for several reasons. The program did not provide instructors any program guidelines—for example, the program did not state that the activities had to be project-based, in any particular format (such as a club), or to follow any type of plan. The site coordinator was unable to describe the main features of the non-academic enrichment activities. Second, the activities were not selected based on a needs assessment. The site coordinator noted that activities were offered as long as there were teachers who were willing to facilitate. This again contrasts with the previous example, in which a needs assessment was conducted, and, in order to address student needs and preferences, the site coordinator hired community members to instruct specialized sessions that could not be facilitated by school-day teachers.

Intentionality Related to Student-Centered Programming

Student-centered programming is a broad term that is used in this chapter, incorporating several types of practices that centers promote to facilitate youth choice, youth voice, and youth leadership opportunities, as well as opportunities for students to interact with one another, and with materials (e.g., tools related to technology, culinary arts, artistic production).

Five of seven elementary centers and two of eight secondary centers in the spring 2012 sample were rated high on student-centered programming. In these centers, respondents described *center-level* approaches to student-centered programming.

One of the elementary centers, for example, exposed students to what the site coordinator referred to as "high quality social interactions, social awareness and social discipline at an early age." These were built into program activities, the site coordinator said. In addition, the center implemented community service points for fifth-grade students who tutored younger youth. In another elementary school, responsibility to others was emphasized, and students were provided opportunities to mentor young students, serve as helpers in setting up the activities, and do other errands. In addition, as an award for participation and good behavior, students were assigned to be "leaders" for Friday activities.

Only two of the high school centers in the site visit sample were rated high on student-centered programming. One of the high schools, which also rated high on intentionality for non-academic enrichment activities, built leadership opportunities into the non-academic activities. The activities were structured as clubs, and within the club structure, students developed products or planned and delivered presentations as a team. In two of the observed activities, students decided among themselves how to showcase their club at an end-of-year presentation. In one of the activities, students worked with a community member to strategize on how to successfully apply for a grant

that would financially support their club activities. An instructor said decisions within the clubs were worked out among students using parliamentary procedures.

Another high school center was rated high on student-centered programming. The center recognized student accomplishments in awards ceremonies. It had also created two ACE "ambassadors" positions, in which designated students assumed responsibility for logistical tasks, such as signing students in, distributing bus passes, and keeping the schedule on track. In addition, within activities, student leaders were selected, and they also supported the management of the activity. Opportunities were somewhat limited, however, because the center's non-academic activities were recreational in nature, and there were no enrichment activities or club structures that might diversify student roles.

A surprising number of the secondary centers did not have clear program-level approaches to student-centered programming. Although observed activities at the centers reflected student-centered approaches, student-centered approaches were not described as an overarching aim of the program by the site coordinators. For example, in one of the centers, the site coordinator indicated that participation in the program in and of itself was indicative of students being leaders. In several high school centers, respondents indicated that any approaches for addressing youth leadership and youth development were up to the instructors.

Summary

Center intentionality was identified as a potential driver of program quality because it clarifies the purpose and focus of activities. Among the 15 ACE programs visited in spring 2012, several factors appeared to influence center intentionality.

- That all elementary programs in the site visit sample—compared with only one of the middle school programs and no high school programs—showed high intentionality with academic enrichment activities suggests that afterschool programming was purposed differently, depending on the grade levels served. The elementary ACE programs had a strong emphasis on academic enrichment activities. These activities were curriculum-driven, linked to school and district objectives, and guided by certain requirements (such as no worksheets). The high school centers typically addressed academic performance through tutoring and not through enrichment activities.

- Center intentionality related to non-academic enrichment activities among the 15 site visit ACE programs was mixed. This may be due to several factors. The most obvious is that some of the centers focused heavily on academic enrichment and considered non-academic enrichment to be a comparatively minor program area. Among the eight ACE programs where center intentionality was high regarding academic enrichment, only four also demonstrated high intentionality related to non-academic enrichment programming. Another factor

may be that among the 15 site visit programs, non-academic enrichment programming was not understood in a consistent way. Only a few of the site coordinators, for example, referred to non-academic enrichment activities as an opportunity for students to learn new skills related to a discipline (such as arts or technology), while others referred to the activities as primarily (and in some cases, only) providing an opportunity for students to socialize, have fun, and/or enjoy a sports activity.

- Intentionality related to student-centered programming was mixed among the 15 centers. Nearly half of the centers appeared to have a deliberate approach to promoting student-centered programming, but most did not. This did not necessarily mean that student-centered approaches were missing from activities, however. Student-centered activities were observed at centers where intentionality at the program level was not high. As one site coordinator said, this is "mainly up to the instructors," and, in some centers, instructors may have been selected because they are comfortable working with students in a way that provides students choice, voice, and leadership opportunities.

Organizational Practices

Organizational practices were examined to determine if and how they varied across the spring 2012 centers. As with intentionality, dimensions associated with organizational practices were described in the program summaries for each site. Each dimension was rated at a high, moderate, or low level, based on evidence of purposeful activities directed toward each dimension.

- *Linkages to the school day* refers to the afterschool program's connection to the school, including awareness of school academic objectives, support from the school, and perceptions from the principal that the afterschool program was aligned with school purposes.

- *Monitoring for improvement* refers to active oversight of the afterschool program and incorporating continuous improvement activities (e.g., monitoring and strategies to improve activities) into the center's operations.

- *Staff collaboration* refers to providing structured opportunities for staff to meet together to discuss students and activities.

- *Staff professional development* refers to providing staff opportunities to learn strategies and content that supports their work in the afterschool program.

- *Site coordinator professional development* refers to the site coordinator learning new strategies and content to support his or her work in managing the program and guiding improvement efforts.

Ratings of the centers on these organizational dimensions were calculated for each center based on interview responses from principals, site coordinators, project directors, and instructors whose sessions were observed. The ratings are summarized in Table 7.

Table 7. Ratings on Organizational Dimensions Among Spring 2012 Site Visit Centers

Center	Linkages to the School Day	Monitoring for Improvement	Staff Collaboration	Staff Professional Development	Site Coordinator Professional Development
Elementary School Centers					
A	High	High	High	High	High
B	High	High	High	High	Moderate
C	High	Moderate	High	High	High
D	High	Moderate	Moderate	Moderate	Mid
E	High	High	High	High	High
F	High	High	High	High	High
G	High	High	Moderate	Moderate	High
Middle School Centers					
H	High	Moderate	High	Moderate	High
I	High	Moderate	Moderate	Moderate	High
High School Centers					
J	Moderate	Moderate	Low	Low	Moderate
K	High	Low	Moderate	Moderate	High
L	Moderate	Low	Moderate	Moderate	Moderate
M	High	Low	Moderate	Moderate	Moderate
N	High	High	Moderate	Moderate	High
O	Moderate	Low	Low	Low	Moderate

Source: Spring 2012 Program Summaries

- The organizational dimensions demonstrated a pattern similar to that of the dimensions associated with intentionality. There was a high degree of uniformity among the elementary centers that were observed. Most ratings were high, and those that were not high were in the moderate range. Only one elementary center had more ratings in the moderate range than the high range.

- For middle and high school centers that were observed, there were more moderate to low ratings, particularly on three dimensions: monitoring for improvement, staff collaboration, and staff professional development. The ratings suggest that the secondary centers in the site visit sample were less centralized in their operations than the elementary centers.

Linkages to the School Day

Close links to the school day were established by almost all of the centers in the spring 2012 sample. This was one approach for improving program quality, not only because such linkages are important for aligning the purposes of afterschool academic enrichment activities and homework help to current learning objectives in the school day, but because close linkages between the afterschool program and the school day likely generate support from campus administrators, which in this sample was reflected in their willingness to help with scheduling and to make space and resources available.

Linkages to the school day were particularly strong within the elementary school centers, which was consistent with their strong emphasis on academic enrichment sessions. The close ties to the school day were evident in several ways, one of which was information transfer from the school-day staff to the afterschool staff. This was reflected, for example, in a center where information was transferred through homework folders that students took to homework help sessions, and there was regular communication among the school administrators, school-day teachers, and the site coordinator. Other ways centers in the sample showed they had fostered close linkages with the school day included data from the school day (e.g., on standardized and benchmark assessments) being made available to the site coordinator, and the alignment of afterschool academic enrichment activities with district and school academic priorities. As a center site coordinator said:

> The district does not dictate to us what we would offer in the afterschool, but we want to align ourselves with the district, and so for example, this year just a few months ago, we had a training on how to write project-based type of curriculum and concept-based curriculum. We take direction from the district and what they're already doing in the daytime to do our own curriculum.

A third type of evidence for strong school and afterschool program ties involved steps taken by the school to support the afterschool program. This was done through providing space, helping with recruitment, and, in some cases, helping with budgets. For example, in one school, the principal allocated part of the school budget to the afterschool program, which had had a recent budget reduction.

Linkages to the school day appeared to have a somewhat different guise in secondary centers than elementary centers. Without a strong emphasis on academic enrichment activities, the observed secondary centers did not generally plan activities around the scope and sequence of courses that were offered during the school day. However, several site coordinators in the secondary school centers spoke of aligning their program with the goals of the school. For example, in one secondary center, where linkages to the school day were assessed as high, the afterschool program was aligned with the campus improvement plan, a process that was done annually. For the 2011–12

school year, the program model at the center was described as "tutorial heavy rather than enrichment heavy to support the students' rigorous school day curriculum." In a high school center where linkages were determined to be in the moderate range, the site coordinator said the principal and he met early in the year to determine academic needs, but made few efforts after this meeting to link the school day and the afterschool program. In several other centers, where site coordinators and principals indicated a supportive relationship between the afterschool center and the school day, interview respondents referred to a very broad alignment to school goals (e.g., wanting students to "participate and excel").

Monitoring for Improvement

Monitoring for improvement is an ongoing process for determining how effectively activities are planned and facilitated, and how activities may be improved to better meet student and school needs. This operational process refers to the monitoring of implementation and the monitoring of results of the centers' efforts—are lesson plans delivered in accordance with the framework and guidelines given sites? Are appropriate instructional approaches used by instructors? Do students appear to be interested in and engaged in lessons? Do students participate? Are there barriers to participation? Are students benefiting from participation?

One of the most notable differences between elementary and secondary centers in the site visit sample was the extent to which the centers engaged in monitoring for improvement. Five of seven elementary school centers and one secondary center (a high school) were rated high on monitoring. The only low ratings on monitoring for improvement were assigned to high school centers. (Four of the six were assigned low ratings on monitoring, as shown in Table 7. The two middle school centers were assigned moderate ratings.)

Nearly all of the elementary centers had distinctive and regular strategies for monitoring and improving activities and programming. Typically, in the elementary centers both the grantee organization's project director and the site coordinator had roles in monitoring programming. They observed activities frequently and provided feedback to instructors on how to improve the activities. The efforts of the elementary school centers and one of the middle school centers suggest that the more a center relies on a curriculum and lesson plans, the greater the efforts are to monitor the delivery of the curriculum. This may be the main reason why monitoring was much higher in the elementary centers in the sample than the secondary centers.

In this sample, the secondary centers' monitoring efforts frequently focused on factors that influenced student attendance: whether activities were what students wanted; and whether logistics, such as transportation and scheduling, supported or impeded student attendance in the afterschool program activities. For example, in a high school center rated high on monitoring for improvement, the ACE program's academic enrichment

activities were in the form of tutorials, led by school-day teachers and aligned with school-day objectives. The non-academic sessions were primarily sports-related. The site coordinator said he monitored program attendance, and that this was his primary strategy for monitoring program quality. On the other hand, for this center, the project director conducted walkthroughs with district curriculum specialists and other district and campus administrators. Decisions based on these walkthroughs included improving the approach of instructors as they led tutorials, and establishing a university-like lounge, which was intended to keep students interested and productive and, thus, further the school's aim to build a "culture of everyone staying after school."

In another high school center, the site coordinator primarily reviewed student attendance to determine whether the non-academic enrichment activities (in the format of clubs) were doing well. When attendance fell off, the site coordinator talked to students about why this had happened, and then addressed any named concerns, such as transportation, schedules, and activities that were not engaging. The project director also reviewed site schedules and observed activities, looking for levels of student engagement and interaction, and ensuring that worksheets were not used.

In two of the high school centers, site coordinators expressed reluctance about monitoring. In one of the two centers, this was due to the site coordinator having no role in selecting instructors; community artists who provided enrichment programming were hired by the grantee organization, and school-day teachers who provided academic programming were selected by the assistant principal. At this center, the site coordinator said he/she had minimal authority over instructors and could not require them to improve the activities they instructed. In another high school center, where services were primarily delivered by high school staff, the site coordinator preferred not to "micromanage" staff. (Site coordinators' willingness to engage staff in improving instructional approaches warrants some investigation in the future. It is consistent with what AIR has observed in high school cultures in general, where teachers often expect a high level of autonomy in developing and implementing lessons during the regular school day.)

Below are several examples of center practices related to monitoring for improvement. The first example comes from an elementary school center, whose practices for monitoring were similar to other elementary centers in the sample. Another example comes from a high school center that was rated high on monitoring for improvement— the only secondary center with a high rating. A third example comes from a high school center where monitoring was assessed as moderate. Finally, a brief description is provided of practices associated with a low rating on monitoring for improvement.

Example 1 – High Level of Monitoring for Improvement: Elementary Center. In this elementary school center, improvement efforts were shared by the project director from the grantee organization and the site coordinator. The project director visited each

afterschool center two to three times a semester. During the visits, he/she met with the principal and conducted observations of activities, checking to see if lesson objectives were outlined on the board, as well as staff preparedness, appropriate use of the curriculum, student engagement, and classroom climate. The site coordinator gathered additional information by surveying parents, students, and teachers to learn their perceptions about activities and what might be improved or changed. According to the site coordinator, the teacher survey inquired about student results: "How are they improving, are they improving academically, are they improving socially, are they raising their hands more or asking more questions?" In addition, the site coordinator frequently, though informally, observed classrooms and provided formative, constructive feedback to staff. The site coordinator said, "I always definitely make sure that it [the feedback session] is one-on-one, and talk to them about things that I'd like to see, or things that they ask me . . . can they maybe change it up, or could they do this as an extra component to further the learning…so I definitely want to give them the opportunity to talk to me as well." Staff reported that the feedback was very beneficial for them.

Example 2 – High Level of Monitoring for Improvement: High School Center. At the only high school center within the observed sample where monitoring and improvement efforts were assessed as high, the project director conducted walkthroughs with district curriculum specialists and other district and campus administrators. These walkthroughs included observations of classroom teachers who were leading tutorials, with observers using the same evaluation forms used during the school-day observations. The project director said, "We use an iPad application for the school-day classes, and have revamped that for the afterschool program. That way, we can collect the same data that is being collected during the school day and there is a clear connection, an alignment of school day and after school if we follow the same model and template." The data were used to drive teacher professional development.

Example 3 – Moderate Level of Monitoring for Improvement: High School Center. At a school that was determined to have a moderate level of monitoring for improvement, a high school center site coordinator and the project director indicated that they conducted observations. The site coordinator conducted informal walk-throughs although was not trained to look for specific practices. (The project director said the grantee organization is working on walk-through protocols that the site coordinator will be trained on.) The site coordinator said student attendance was the primary criterion to determine quality, and if attendance dropped, the site coordinator talked with students to determine why and then worked to correct problems. Solutions included working with the instructors to make sessions more engaging, providing later transportation, adjusting schedules, and replacing activities with others students found more appealing.

Other Examples – Low Level of Monitoring for Improvement. Four of the high school centers in the site visit sample appeared to minimally monitor for improvement. In these centers, logistics were typically a greater concern than the quality of sessions. In one site, the project director monitored activities and data that were in the tracking system, conducted site visits, and observed activities. At this center, the site coordinator mainly monitored activities by looking at student report cards and learning from instructors how students were doing and whether (or how) they were improving either academically or behaviorally. Follow-up improvement strategies were not mentioned. In another high school center, staff meetings mainly focused on what activities were offered and planned, supplies, and calendars. In another high school center, the site coordinator said he or she had no control over instructors with the agency the school contracts with, and that if activities were not well attended, they were dropped and different activities were put in their place. There was no mention of improving the quality of activities.

Staff Collaboration and Staff Development

Another approach for supporting the quality of activities is staff collaboration and staff development. These were examined because they provide opportunities for staff to work and learn together as they improve their instructional practice and the design and delivery of activities. Staff collaboration refers to staff working together in scheduled sessions to systematically address programming: what works, what does not work, for what reasons, and what can be done about it. Staff development refers to formal opportunities for staff to learn together in focused sessions facilitated by an expert— someone with more knowledge and a different perspective from the staff. Not only may staff collaboration and staff development improve program quality, but, as will be evident in the following discussion, these approaches may also establish a consistent programmatic approach for developing and implementing program activities.

Staff collaboration was rated high when staff met frequently and in scheduled sessions, and when designing and improving the activities (not just deciding which activities to offer) were a function of collaboration. Moderate range ratings were assigned when collaboration was evident but designing and improving sessions was not a focus of collaboration; when collaboration was informal and depended on instructors finding time on their own to collaborate; or when collaboration among instructors was low, but instructors did meet with the site coordinator to improve sessions. Low ratings were assigned when there were few, if any, opportunities to collaborate.

Five of the elementary centers in the sample were rated high on both staff collaboration and staff development, and only two were rated in the moderate range on dimensions. This was different from the secondary centers in the site visit sample, where one of the eight centers was rated in the high range on collaboration and none were rated in the

high range for staff development. With one exception, the centers in this sample were assigned the same rating for staff development as staff collaboration, suggesting that the two organizational processes are closely related.

Examples – High Level of Staff Collaboration and Staff Development: Elementary and Middle School Center. The elementary centers and one of the middle school centers in the sample showed higher levels of collaboration. In an elementary center, staff met every Monday to discuss what was occurring in the program and to work on lesson plans. Most of the afterschool activities were facilitated by two instructors who worked together to design the lessons. The site coordinator said collaboration had prevented staff from feeling overwhelmed and had also made lesson planning easier. In a middle school center, the instructors met weekly to discuss concerns, progress, and activities for the week. Instructors frequently collaborated to modify centrally developed lesson plans to better meet the needs of their students. They also collaborated to share effective approaches for addressing particular issues they encountered in their sessions.

Examples – Low Level of Staff Collaboration: High School Centers. The level of collaboration and staff development described in the prior example contrasts with that of some centers in the sample, rated low, where staff had fewer opportunities to interact. In one high school center, the site coordinator and instructors indicated that they talked with one another daily about "anything that is going on," but that this was mostly done when staff picked up and dropped off their daily sign-in sheet—not in a formal group meeting. While program staff were responsible for planning the activities they facilitated (both academic and non-academic), they rarely communicated with each other about the program or students in the program. In another high school center, frontline staff met together three times a year. It was difficult to hold more frequent meetings because instructors, most of them certified teachers, had other obligations. The staff members communicated with the site coordinator mainly by email and text messaging. The site coordinator said, "Everybody is kind of individualized because nothing really overlaps." Some staff members collaborated with community partners but did this on their own.

The descriptions of staff development practices also suggest that staff development is closely related to program monitoring for continuous improvement. Those centers with strong staff development based much of their training on what site coordinators and project directors observed in the sessions and/or what staff told them about their professional development needs.

Example – High Level of Monitoring for Improvement and Staff Development: Elementary Center. The elementary center in this example aligned staff development with information learned through monitoring. In this center, the site coordinator identified staff needs for professional development based on classroom observations and through discussions with the youth leaders. Two main areas for improvement were identified

and addressed this year: parent interactions and classroom management. With the district offering a wide array of training options for afterschool staff, some required and some optional, the site coordinator provided suggestions to the youth leaders about the training that would be most beneficial. The instructors who were interviewed reported that the training they attended helped them identify and apply useful resources and approaches. The site coordinator said, "We are training constantly, in the evenings, on a Saturday, and on the early release days. There are always opportunities."

Example – Moderate Level of Staff Development: High School Center. In the following high school center, staff development was available for youth workers, but not for the school-day teachers who facilitated the majority of the sessions. Professional learning opportunities were available every Friday for two or three hours to staff who were not certified teachers in the school. The sessions were not mandatory, but the site coordinator hoped to change this policy in the future. The site coordinator said the staff development sessions were focused on steering the staff away from "babysitting" and the idea that "they really don't have to tie in academic unless it's a tutoring," which she described as "a big problem." The school-day teachers, who made up the majority of the program staff, did not receive any professional learning related to afterschool instruction. Although the school-day teaching staff were able to link academics into the afterschool program, the site coordinator said they did not know how to make the activities fun, and needed training in this area.

Example– Moderate Level of Staff Development: Elementary Center. This example, and the example above, show that youth workers who were not employed by the schools were provided more professional development than the certified teachers who planned and facilitated activities. In this elementary center, most school-day teachers received professional development through the school district and were not offered additional professional development by the afterschool program, although some of the teachers sought professional development related to their activity area. Program instructors who were affiliated with outside vendors received training from their organizations.

Professional Development for Site Coordinators

The site coordinators received fairly consistent professional development. Nine of the site coordinators from the spring 2012 sample received a high level of professional development related to their responsibilities, and six received a moderate level of professional development. Much of the professional development was provided at the regional level, and nearly all site coordinators indicated it was useful, particularly with respect to program and staff management. Two site coordinators had received professional development on the Youth PQA, and one site coordinator mentioned the Colors training, which identifies different personality types and ways to work with people of different types. Among the more positive assessments of the training for site

coordinators are the two following examples.

Example 1 – High Level of Professional Development for Site Coordinators: High School Center. In this example, the high school center coordinator participated in monthly mandatory meetings provided by the grantee organization. The trainings were relevant to the school district conditions and local needs. Some of the sessions featured local agencies, such as an agency for the homeless and the agency that addressed child abuse and neglect. Some sessions focused on academic program components, and others addressed program management. The site coordinator said the training sessions were helpful. Some of the topics were closely aligned with the district's systems and policies, and others made her much more aware of social issues encountered by students, such as hunger, and child abuse and neglect.

Example 2 – High Level of Professional Development: Elementary Center. In another center, the site coordinator attended United Way professional development sessions and passed on what he had learned to his staff. He was also certified as a Youth PQA assessor. He typically attended two afterschool conferences a year. In addition, the site coordinator received weekly professional development through his grantee network, usually delivered by the project director. Topics covered included project management and effective strategies for addressing bullying. The project director also assigned different topics for the site coordinators to research and then present to their peers.

Community Connections

The last operational component examined is the connection the center has made to the community, as evidenced by the extent to which the community has contributed to the center. Community connections can be particularly helpful to centers and contribute to program quality because they can increase programming options and resources for both students and their families. Of the 15 centers, 11 were assessed as having a high level of community connections; and in only one center were community connections assessed as low. The two examples presented describe how the center's connections to the community enhanced programming options and resources.

Example 1 – High Level of Community Connections: Elementary Center. In an elementary center, the site coordinator and the project director both named developing relationships with community organizations as one of their major responsibilities. The project director established relationships at a regional level and connected resources from the community to different centers that needed support. The site coordinator worked with the public library to connect parents to adult learning opportunities (e.g., computer classes, citizenship classes), and also developed connections with area homeless shelters and child abuse departments to assist families with more serious needs. The site coordinator had also established a relationship with a local non-profit

advocacy group so that students could work on a community service project. In addition, the center connected students to institutions (the art museum and public libraries) so students became aware of the resources available to them. The site coordinator said that before a field trip to the library, many students did not realize that they could get a library card and check out books for free, nor were they aware that the library offered programs for both youth and adults.

Example 2 – High Level of Community Connections: High School Center. In one of the high school centers, the instructors made extensive efforts to partner with community members. One of the teachers partnered with a local production company, whose staff mentored the students in a film club. Another staff member teamed up with empowerment groups at the local university. The partnerships greatly improved opportunities for students in the afterschool program. One of the instructors said that, as a result of the partnership, students screened their films in local movie theaters, and students spoke in public to introduce the films they produced in public venues. Several students became interns with the agencies and businesses the center partnered with.

Summary

This chapter addressed research question 3: "What organizational processes are found to be drivers of instructional /point of service quality at high performing centers?" In particular, the chapter examined center intentionality with respect to programming, and organizational practices that have the potential to develop and deliver high quality activities and improve those activities on a regular basis. Several findings that were presented in this chapter show sharp differences in secondary (particularly high school) and elementary programming and operations. The findings that are summarized here will highlight those differences when relevant.

- Among the 15 ACE programs, center intentionality with respect to academic enrichment programming showed a strong focus on meeting academic objectives that were emphasized during the school-day core classes. The centers had developed appropriate curriculum, which included interactive learning, a variety of learning experiences, and, when appropriate, hands-on learning. Because the elementary centers in the sample had more of a focus on academic enrichment activities than secondary centers, center intentionality with respect to academic enrichment programming was more evident in elementary centers.

- The examination of center intentionality showed that there is not a shared understanding of what non-academic enrichment programming is. In several centers in the site visit sample, the 'enrichment' aspect of the non-academic activities was missing. Enrichment-type activities teach students knowledge and skills related to a specific discipline, such as film-making or choreography.

However, numerous respondents referred to these activities as opportunities to relax, play games, and be physically active. Knowledge and skills specific to a discipline were not described. In these centers, intentionality with respect to non-academic enrichment was not high because the purposes of the activities were not centered on learning objectives.

- Practices to monitor for improvement (that is, to improve the quality of activities) were evident in the majority of the centers in the site visit sample. Particularly strong were the centers that systematically approached monitoring of academic enrichment activities and modifying curricula and instructional approaches to better engage students and meet learning objectives. To an extent, this was more obvious among the elementary centers in the sample. In secondary centers, where monitoring was assessed as high, attendance was sometimes the foremost factor in monitoring and making improvements in programming. Site coordinators and, in some cases, project directors considered student attendance and student motivation to participate—that is, were students interested in the activities? Were the sessions themselves appealing? Were there barriers to attendance that needed to be addressed? Responses to monitoring of this type included implementing different types of activities, particularly non-academic enrichment activities, based on student interest, and working with instructors to better align instructional methods with students' developmental needs. This type of monitoring seems appropriate in high schools where students have a high level of choice regarding participation.

- When centers had established close ties with the school, it was reflected in information transfers, availability of space, and administrative support. The elementary centers generally aligned their academic enrichment activities to the scope and sequence of the day-school classes. The secondary centers, particularly high schools, more broadly aligned activities to the general goals of the school.

- Staff development and staff collaboration were other means of supporting program quality. Among the 15 centers in the 2012 site visit sample, staff development and opportunities for collaboration, particularly collaboration in scheduled sessions, were more available to instructors who were not certified teachers than to those who were teachers during the school day. The latter had little time to meet formally, though they did have some opportunities for sharing information on an informal basis. The scheduling constraints, a site coordinator said, made it difficult to improve instructional methods of staff in order to make them more engaging for youth. In this group of ACE programs, the elementary centers were more reliant on non-certified teachers to instruct sessions than secondary centers, where activities were frequently instructed by certified

teachers. This is particularly the case when non-certified teachers are associated with the grantee organization, rather than when they are community members. Thus, in this sample of programs, it was elementary centers more than secondary centers that had opportunities for both staff development and collaboration for at least a majority of instructors. Providing opportunities for staff development and collaboration for certified teachers may be a persistent challenge for many afterschool programs. Sometimes, this is addressed through the professional development of site coordinators, which was extensive for the majority of the centers in the sample.

- Community connections were important to many of the centers, most of which had developed partnerships with organizations and agencies that resulted in expanded opportunities for programming, youth, and parents.

The next chapter continues the exploration of center quality by showing its influence on student participation and student academic and behavioral outcomes among the random sample of Cycle 5 and Cycle 6 centers where site visits were conducted in spring 2011. Participation will be considered in terms of intensity (how many days of participation), and duration (over how many days did participation last). Outcomes include academic achievement and behavioral outcomes.

Chapter 4
The Influence of Center Quality on Student Participation and Outcomes

In 2011–12, the two primary research objectives of the evaluation of ACE programs were to understand how well centers were implementing *quality* programming in terms of research-supported practices and approaches, and what impact *participation* in ACE-funded activities had on student academic outcomes.

The 2010–11 Interim Report[26] detailed findings showing that the span of quality associated with the Cycle 5 and Cycle 6 programs was quite broad—that the centers differed in the extent to which they adopted practices that supported academic skill-building among participating youth, as well as practices that supported youth development. For 21st CCLC programs nationally, practices that support both academic skill-building and youth development are at the heart of an academic enrichment program model that focuses on academic objectives while providing youth opportunities for choice, voice, and leadership. In the Interim Report, one of the findings was that there was considerable variation among the ACE programs with respect to the use of practices likely to support academic skill-building: these practices include intentionally designing programs, establishing linkages with the school day, and using student data to inform the design and delivery of programming. Findings noted that stronger linkages to the school day were evident among centers that were mostly staffed by school-day teachers, unless the center served high school students. Less intentionality in designing programs was evident among ACE programs serving high school students than those serving middle and elementary school students. Use of student data was generally low across centers, though somewhat higher in the centers staffed mainly by school-day teachers.

Practices that support youth development also varied across centers. For example, centers serving students in secondary schools were more likely than other centers to provide youth opportunities for ownership in the program (e.g., choosing the type of activities that would be offered, having some voice in how those activities were managed). Practices associated with youth ownership are more viable with older students who may have developed the skills to effectively exert ownership.

Because findings from the Interim Report demonstrated that ACE-funded programs varied in implementing quality-related practices, it may be that the programs varied as well in their effectiveness in recruiting and retaining students and achieving desired youth outcomes, e.g., improved academic achievement and fewer disciplinary incidents.

---

[26] http://www.tea.state.tx.us/index4.aspx?id=2908&menu_id=949

In this section of the report, we explore the following questions related to the relationship between program quality, youth participation, and student outcomes.

- What is the relationship between the characteristics of individual youth, center quality, and other center characteristics and levels of student participation in ACE programming?

- Does the impact on student outcomes vary by relevant ACE program characteristics, including center quality?

These two questions are representative of both TEA goals and objectives for the ACE program and pressing questions currently before the afterschool field. As noted by Granger (2008), much of the research shows afterschool programs have a mixed impact on students' academic and behavioral outcomes. For example, three noteworthy meta-analyses of studies conducted in the field of afterschool research (Durlak & Weissberg, 2007; Lauer et al., 2006; Zief, Lauver, & Maynard, 2006) explored the impact of afterschool programs on student achievement and behavioral outcomes. In each of these three meta-analyses, the authors found that in the *majority* of the studies reviewed, students in the afterschool program did not have better outcomes than a comparison group of non-participating students. However, when completing a meta-analysis of the program impacts reported in studies included in each review, both Durlak and Weissberg (2007) and Lauer et al. (2006) found average *positive* effects in both academic and non-academic outcomes, indicating that a smaller number of highly effective programs in each review had driven the positive effects across all programs. That is, positive outcomes across multiple programs were likely due to the effectiveness of a small number of individual, *high quality* programs. Durlak and Weissberg showed that programs with large, positive effects on student outcomes were characterized by the adoption of sequential approaches to personal and social skill development: skills targeted for development in one session were based on skills cultivated in prior sessions. In addition, instructors in the high quality programs facilitated active forms of learning, focused on developing specific skills, and were explicit in telling youth the skills the activity aimed to introduce, teach, and/or cultivate. The work conducted by Durlak and Weissberg made an important contribution to identifying approaches that are likely to lead to desired youth outcomes. Further research is needed to determine what programmatic features and characteristics drive program outcomes, and to develop effective quality improvement systems and related interventions to support the cultivation of these attributes (Granger et al., 2007).

Approach for Developing Quality Profiles and Variables Used in the Analyses

A primary goal of this chapter is to take a theoretical model for quality service delivery in ACE programs and explore evidence suggestive of a link between high quality

programs based on this model and the achievement of desired youth outcomes. This section explains how the analyses presented were conducted by describing the data on quality used to identify the quality of programming at the centers; the development of quality profiles; and finally the other center- and youth-level variables that were included in the analyses.

Data on Quality

The data for the analyses presented in this chapter are from the 40 Cycle 5 and 6 ACE programs visited in the spring of 2011 and the 40 Cycle 5 and 6 programs visited in the fall of 2011. Data included observation data, for which scoring related to quality was applied using three observation protocols—the PQA, the OCE, and the APT-O; and staff surveys, which included constructs related to quality, that were administered to staff at each of the centers where site visits were conducted. (Additional information about the observation protocols is provided in Appendix B.)

Data associated with centers visited in 2011 were used to both (a) create quality profile types meant to classify centers into different quality tiers, and (b) construct variables that could be used in a series of analyses designed to assess the relationship between program quality and outcomes. Data from all 80 centers visited in 2011 were used to develop center quality profiles. When conducting analyses designed to answer the specific research questions on the relationship between program quality and participation and student outcomes, however, only data from the 40 centers in the spring 2011 sample were included. This was done for two reasons:

- The spring 2011 sample was a random sample of all Cycle 5 and 6 centers operating during the 2010–11 school year. As a result, variation across programs in the level of quality was anticipated. This was not the case with the fall 2011 sample. The fall sample was selected based on indicators suggesting the centers were higher quality. Thus, a high degree of variation with respect to quality was not anticipated, and this would likely make it more difficult to identify significant relationships between program quality and outcomes.

- The participation and student outcome data available to support the analyses presented in this chapter were associated with the 2010–11 school year. The evaluation team believed it was important to consider outcomes associated with the same time period in which data used to measure program quality were collected. In prior and subsequent years, some of the centers may have changed their program design and delivery, which would make any measures of program quality less meaningful.

From the 40 spring 2011 centers, data included PQA scores from 157 observed activities; OCE scores from 110 observed activities; and APT-O scores from 85 activities. In addition, 465 staff surveys from staff at the 40 centers were completed;

between one and 26 surveys were received from each center, with an average of 12 surveys per center. However, to enhance the stability and power of quality estimates derived from observations and staff surveys, data obtained from the full 80 centers represented in the spring and fall 2011 samples were analyzed together utilizing Rasch analysis techniques. Additional information about each of these data sources can be found in Appendix B.

Developing Quality Profiles

One of the goals of the spring 2011 site visits was to obtain data on the adoption of quality practices that support both academic skill-building and youth development from a variety of sources. These data could then be triangulated to fashion a comprehensive portrait of the level of program quality at a given center. Data collected from activity observations provided an objective, external assessment of quality at the point of service. However, the capacity to generalize observation results to formulate an overall estimate of quality at a given center was weakened by the fact that a minority of staff were observed across only two program days of the site's school year program. Although the staff survey was limited due to the self-report nature of the items, the majority or all of the staff completed surveys. In addition, the surveys were non-event-based, and items addressing quality were not time sensitive, as was the case with observations.

Both survey and observational data were analyzed using Rasch analysis techniques (see Appendix C for a fuller description of these approaches), which resulted in a 0 to 100 scale score for the activity (in the case of observations), or for the respondent (in the case of the staff survey). Higher scores indicated a higher level of quality. For the observation measures, total instrument scores were averaged across each of the activities observed to derive a center-level quality estimate, shown in Table 8. A similar approach was used with the staff surveys, where individual staff responses were averaged to develop a center-level quality estimate. In the case of the staff surveys, center-level quality estimates were derived for four different subscales, shown in Table 8.

Table 8. Summary of Measures by Construct Used to Create Quality Profiles

Measure	Practices Supportive of Academic Skill Building	Practices Supportive of Youth Development
Point of Service Quality		
PQA		Total Score
OCE		Total Score
APT-O/PQA Academic Climate	Total Score	
Non-Event Measures of Quality		
Staff Survey	• Intentionality in Program Design Scale Score • Practices Supportive of Academic Skill-Building Scale Score	• Practices Supportive of Positive Youth Development • Youth Ownership Scale Score

Although each subscale identified in Table 8 was considered individually in the Interim Report, the evaluation team decided it would be more useful to pool the subscales into a series of quality types or profiles and then assign each center to a quality type based on its score. Two sets of hierarchical cluster analyses were run to assign each center to a quality profile type, one employing the observation data and one using data associated with each of the staff survey scales under consideration.[27] The goal of these analyses was to develop between two and five quality profile types, ranging from low overall quality to high overall quality, and then classify each of the spring 2011 centers as one of the types. Membership in a given profile type would then be used as a predictor in models constructed to assess if a relationship existed between centers assigned to a quality profile type on the one hand, and student participation and student outcomes on the other.

First, a hierarchical cluster analysis was conducted, using scores from the PQA, the OCE, and the APT-O, to classify centers into one of five quality profile types. This analysis was then re-run four more times, yielding results for four, three, and two quality profile types. A similar process was employed with each of the four staff survey scales presented in Table 8. Ultimately, the number of quality profiles selected was based on how well the categories differentiated programs into homogenous categories that made

---

[27] Observation and survey data were analyzed separately because it was determined that conceptually it made more sense to have one profile assignment based on external, point-in-time measures and another predicated on a broader spectrum of staff responding to questions that were less occasion dependent. In addition, cluster analysis is typically employed to combine cases (or, in this case, centers) into groups using a series of variables as criteria to determine the degree of similarity between individual cases, and is particularly well suited when there is a desire to classify a large number of cases into a smaller domain of discrete groupings.

good interpretative sense. For the observation-based profiles, the 4-quality profile type solution was found to meet these criteria (see Figure 2). The four types are:

- *High POS Quality.* Centers assigned to this cluster were characterized by high levels of average program quality at the point of service (POS) on each of the three observation protocols employed. It was hypothesized that centers in this group would be the most closely associated with high levels of student participation and greater program outcomes. A total of 21 of the 80 (26%) centers examined were classified in this profile type, and 10 of the 40 (25%) centers visited in the spring of 2011 were classified in this profile type.

- *Low POS Quality.* Centers assigned to this cluster were basically the inverse of centers assigned to the *high POS quality* cluster, with low average scores across each of the observation measures. A total of 12 of the 80 (15%) centers examined were classified in this profile type, as were 4 of the 40 (10%) visited in the spring of 2011.

- *High APT-O/Academic Climate.* Centers assigned to this cluster scored reasonably well on the APT/PQA Academic Climate measure, on average, but not quite as well on either the OCE or the PQA. Generally, centers in this cluster were of moderate quality, performing better on measures related to supporting academic skill-building than on other measures. A total of 28 of the 80 (35%) centers examined were classified in this profile type, and 15 of the 40 (38%) centers visited in the spring of 2011 were classified in this profile type.

- *High OCE.* Centers in this cluster scored quite high on the OCE, on average, but scored relatively poorly in relation to the PQA and APT-O/Academic Climate measures. Generally, while activities provided by these centers were engaging for youth, there was little evidence that staff engaged in practices intended to support academic skill-building or youth development. A total of 19 of the 80 (24%) centers examined were classified in this profile type, as were 11 of the 40 (28%) centers visited in the spring of 2011.

Figure 2. Quality Profiles Based on Observation Data, Spring and Fall 2011
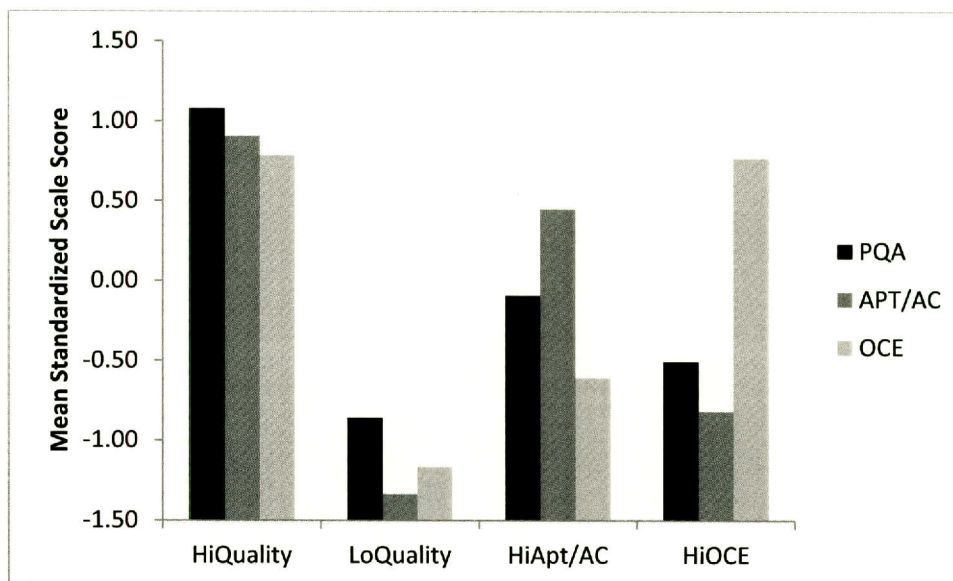
Source: n=238 activities observed at 80 programs

For the staff survey-based profiles, the 2-quality profile type solution was found to be the most interpretable (see Figure 3). It was hypothesized that the staff survey would be more effective in identifying lower quality programs than distinguishing among programs at the higher end of the quality continuum. This hypothesis was derived from the possibility that some respondents might answer survey items in a socially desirable way, thereby introducing more error into quality estimates at the higher end of each survey subscale.

- *Higher Reported Quality.* Centers in this cluster had higher average scale scores values across each of the scales considered: intentionality in program design (*Design*); practices to support academic skill-building (*ASB*); practices supportive of youth development (*YD*); and youth ownership (*Own*). A total of 47 of the 78 (60%) centers examined were classified in this profile type, and 23 of the 40 (58%) centers visited in the spring of 2011 were classified in this profile type.

- *Lower Reported Quality.* Centers in this cluster had lower average scale scores values across each of the four staff survey subscales. A total of 31 of the 78 (40%) centers examined were classified in this cluster, including 17 of the 40 (43%) centers visited in the spring of 2011.

Figure 3. Quality Profiles Based on Staff Survey Data, Spring and Fall 2011
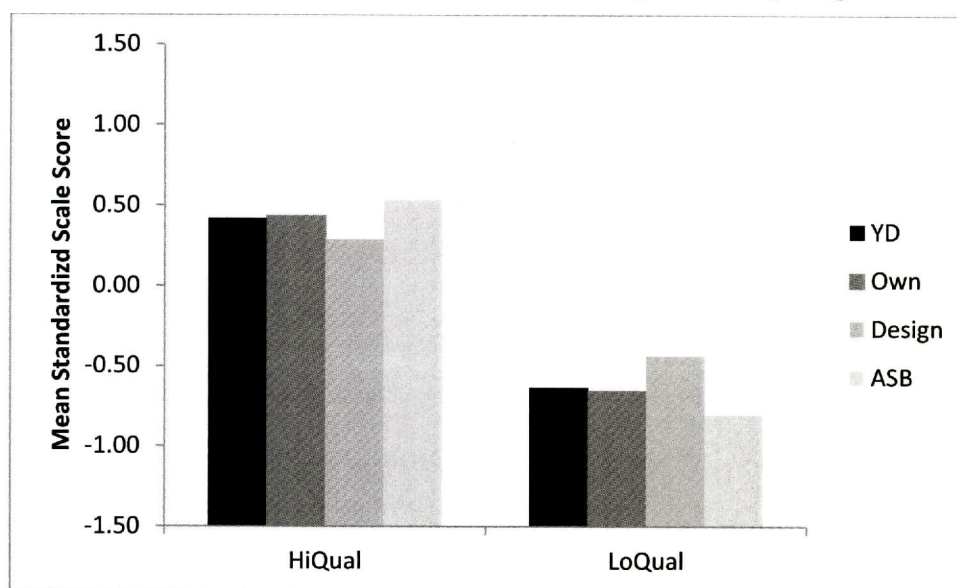
Source: *n*=790 staff surveys collected from 78 programs

A comparison of center membership in quality profiles derived from observations and the staff surveys demonstrated a moderately significant, positive correlation, indicating that some consistent signal about the level of program quality at each site was being detected through the profiles that were constructed ($p < .10$, Chi-Square = 6.851, df = 3).

Other Center and Youth Variables

In addition to center quality, a series of additional center- and youth-level variables were included in the analyses as predictors for assessing the relationship between program quality and youth outcomes. Most of these variables were selected based on findings from previous evaluations that demonstrated a relationship between center- or student-level variables and student outcomes (Naftzger et al., 2012; Naftzger, Vinson, & Swanlund, 2012; Naftzger, Vinson, Manzeske, & Gibbs, 2011; Naftzger, Vinson, & Swanlund, 2011). Some of the variables were also used as outcomes in analyses examining the relationship between program quality and student participation in the program; these are the total number of hours a student participated in ACE programming during the 2010–11 school year, and the total number of calendar days between the student's first and last day of participation.

Three center-level variables were included in the analyses. These were:

- **School-Based Status.** This refers to whether or not the center was associated with a school-based grantee. It was hypothesized that centers associated with school-based grantees would find it easier to gain access to student achievement

data that would inform the design and delivery of programming; and that they would have established linkages with the school (and thus, would be more informed about the curriculum and lesson objectives, and communicating more frequently with school-day teachers). Analyses undertaken in relation to the Interim Report demonstrated that activities observed in centers associated with school-based grantees were more likely to score higher on the *interaction* and *engagement* scales of the PQA.

- **Center Maturity.** This refers to how many years the center had been an ACE grantee. Centers in their first year of operation as an ACE program were not expected to be as high functioning as more mature programs. Although results presented in the Interim Report did *not* support this hypothesis, evidence supporting the hypothesis was found in other 21st CCLC statewide evaluations conducted by the evaluation team, with more mature afterschool programs in elementary schools showing stronger academic outcomes than less mature programs (Naftzger, Vinson, & Swanlund, 2012).

  - **Staffing Model.** This refers to whether or not the center was staffed mostly by school-day teachers. Again, it was anticipated that programs staffed mostly by school-day teachers would be more likely to adopt practices related to academic skill-building because they would have knowledge of students needs, familiarity with school curricula, and greater access to and familiarity with student achievement data. This hypothesis was supported by results that were presented in the Interim Report.

Five youth-level variables were developed. Except for the first variable, grade level, the variables refer to patterns related to student participation in the ACE programming.

- **Grade Level.** The grade level of the student was categorized as elementary, middle, or high. Elementary programs have been shown to have higher rates of attendance in ACE programming than either middle or high school programs (Naftzger et al., 2012; Naftzger, Vinson, & Swanlund, 2012; Naftzger, Vinson, Manzeske, & Gibbs, 2011; Naftzger, Vinson, & Swanlund, 2011). In addition, in other statewide evaluations conducted by the evaluation team, measures of program quality similar to those in this study were found to be more predictive of student outcomes for elementary students than secondary students (Naftzger, Vinson, & Swanlund, 2012).

- **Hours of Participation.** A second youth-level variable examined was the student's number of hours of participation in ACE programming in the 2010–11 school year. In analyses that assessed the relationship between program quality and participation (summarized later in this chapter), this variable was used as an outcome. In analyses exploring the relationship between program quality and

student outcomes, it served as a predictor, given the hypothesis that greater program participation would be positively associated with program outcomes.

- **Duration of Participation.** Duration refers to the span of participation: the number of calendar days between the student's first day of participation in the ACE programming (in 2010–11) and his or her last day of participation. In analyses that assessed the relationship between program quality and participation (summarized later in this chapter), this variable was used as an outcome, allowing the evaluation team to explore whether program quality maintained student participation (and engagement) in programming for a period of time during the school year.

- **Activity Type.** Another youth-level variable is activity type, which is intended to measure the diversity of activities a student participated in during the 2010–11 school year. The variable indicates the number of different activity categories in which a student spent at least 10 percent of his or her total participation hours. Having an opportunity to participate in a variety of activity offerings is one common element of program quality represented in many different quality frameworks for afterschool and is considered to be conducive to student engagement (Little, 2007). This variable was not examined in the Interim Report.

- **Alignment with Academic Activities.** This variable indicates a student's degree of alignment with either a *high academic enrichment* or *high homework help* activity profile. As grantees populate the TX21st database with data, they record student attendance in specific activities at each of their ACE programs. Information is also provided about the following types of activities students attend. These include the following:

  - Academic enrichment learning program
  - Recreational activity
  - Homework help
  - Supplemental Education Services tutoring
  - Activity to promote youth leadership
  - Expanded library service hours
  - Drug/violence prevention, counseling, or character education
  - Career/job training
  - Promotion of family literacy
  - Mentoring
  - Community service/service learning

- Other (e.g., activities involving computers and technology, life skills, nutrition, etc.)

Variables indicating the total proportion of time a student spent participating in each type of activity were used to identify two activity profile types: a *high academic enrichment* profile, indicating students who spent most of their time attending academic enrichment offerings and little time in recreation activities; and a *high homework help* profile, indicating students who spent the majority of their time at ACE programs working on homework.

It is important to note that activity profiles do *not* represent actual students served by the ACE program, but serve as markers that can be used to determine if a student more closely resembles one type of student profile or another. Thus, the two primary student activity profiles allowed students to be identified by the profile they most resemble. The activity profiles are useful because they help determine whether students within each profile type are associated with high levels of program participation and positive student achievement and behavioral outcomes, allowing a comparison of outcomes by profile.

Summary

This section described how quality profiles were developed for each center, including those developed from observation data and from staff survey data. Together, the profiles represent quality at different levels (higher and lower levels) and types (e.g., related to academic programming, and student engagement). In addition, the section described the center- and youth-level variables that were used to explore the research question on the relationship of program quality to student outcomes. The remainder of the chapter will first describe center quality and its relationship to student participation; then the relationship of participation to center and student characteristics; and finally, the relationship of center quality to student academic and behavioral outcomes. The chapter concludes with a summary statement on what the analyses showed about the relationships that were explored, and the implications of this for afterschool practice.

The Relationship of Center Quality to Participation

In this section of the report, results are presented related to the following research question: "What is the relationship between the characteristics of individual youth, center quality, and other center characteristics and levels of student participation in ACE programming?" The hypothesis was that higher program quality (measured using observation and staff survey data from the 40 ACE-funded centers visited in the spring of 2011) would be associated with higher levels of student participation. Both the hours of programming and the duration of attendance were used as participation outcomes in two multilevel models, one model for each outcome. (See Appendix D for more details.) Participation data were available for 10,381 students attending the spring 2011 site visit

centers. The average student attended 114 hours of ACE-funded programming over a span of 176 days.

The center- and youth-level variables were included as predictors in each analytic model. As shown in Table 9, in terms of student characteristics, students were fairly uniformly distributed across each of the three grade levels, with 40% in elementary school, 31% in middle school, and 30% in high school. On average, a student participated in just over two different types of activity categories (e.g., academic enrichment plus recreation; tutoring plus service learning, etc.), depicted by the mean of 2.19.

Table 9. Student Characteristics for Centers Visited in Spring 2011

	2010–11 ACE Participants Attending Site Visit Centers (N=10,381)
	Percentage
Grade Level: Elementary	39.74%
Grade Level: Middle	30.69%
Grade Level: High	29.57%
	Mean
Activity Category Diversity	2.19

Source: TX21st

Center-level predictors are depicted in Table 10. Centers represented in these analyses tended to be school-based (80%); in their second year of operation (53%); and they tended to employ an approach to staffing that was not primarily made up of school-day teachers (65%). The 40 centers were also assigned fairly evenly to different quality clusters. With one exception, each cluster (within the 4-cluster quality types and the 2-cluster quality types) was comprised of at least 25% of the centers.

Table 10. Center Characteristics for Centers Visited in Spring 2011

	2011 ACE Site Visit Centers (N=40)
School-Based Status	Percentage
School-Based	80.00%
Non-school-Based	20.00%
Maturity	Percentage
First year of grant	20.00%
Second year of grant	52.50%
Third year of grant	27.50%
Staffing Model	Percentage
Staffed mostly by school-day teachers	35.00%
Staffed mostly by other types of staff	65.00%
Program Quality Measures	Percentage
Observation: High POS Quality Cluster	25.00%
Observation: Low POS Quality Cluster	10.00%
Observation: High APT-O/Academic Climate Cluster	37.50%
Observation: High OCE Cluster	27.50%
Staff Survey: Higher Reported Quality Cluster	57.50%
Staff Survey: Lower Reported Quality Cluster	42.50%

Source: TX21st, Observations, and Staff Survey

The Relationship of Participation to Center and Student Characteristics

Multilevel models were used to explore the relationship between participation outcomes among students enrolled in ACE programs in 2010–11 that were visited in spring 2011, and the center and student characteristics shown in Tables 9 and 10. Two separate models were run, one model in which the outcome was the total *number of hours* of 2010–11 ACE programming a student participated in, and one model in which the outcome was the *number of calendar days* between the student's first and last day of participation in 2010–11 ACE programming. (See Appendix D for a more detailed description of this approach.)

The reader should keep in mind that the findings described in this section of the report are purely descriptive in nature and do not in any way imply that a given program or student characteristic was found to be *causally* related to a given participation outcome.

The results should be considered exploratory, requiring further confirmation through a more robust design that better controls for issues of selection bias.

Total Number of Hours of ACE Programming

Several center- and youth-level variables were found to be related to the number of hours of ACE programming a student participated in during the 2010–11 school year. When considering center quality, as shown in Table 11, center membership in the *lower reported quality* cluster (derived from the staff survey) was significantly and negatively associated with the number of hours of participation. Students attending centers in the *lower reported quality* cluster participated in fewer hours of ACE-funded programming than students associated with centers in the *higher reported quality* cluster. As noted earlier, centers assigned to the *lower reported quality* cluster were less likely to adopt the following quality-related practices: intentionality in activity and session design; practices supportive of academic skill-building (including linkages to the school day and use of data on student academic achievement to inform programming); practices supportive of positive youth development; and opportunities for youth ownership.

In addition, students attending centers falling in the *high OCE* cluster were also found to attend fewer hours of ACE programming during the 2010–11 school year than students enrolled in centers classified in the *high APT-O/Academic Climate* cluster, a moderate quality category. While activities provided by centers in the *high OCE* cluster were engaging for youth, there was little evidence of staff engaging in practices that supported academic skill-building or youth development. Practices measured via the PQA and APT-O/Academic Climate scales were *less* prevalent in activities associated with centers in the *high OCE* cluster. That is, the centers were less likely to create a supportive environment, foster positive interactions between youth and the activity leader, and create engaging learning experiences (as determined by the PQA). They were also less likely to encourage students to practice content-specific skills in reading, written and verbal communication, mathematical communication and reasoning, and mathematical problem-solving (as determined by the APT-O Academic Climate scale).

On the other hand, as shown in Table 11, there were no significant differences in hours of participation between students enrolled in centers assigned to the *high POS quality* cluster and those in centers assigned to the *low POS quality* cluster, when compared to students enrolled in the *high APT-O/Academic Climate* cluster. This result was *not* expected. The hypothesis of the evaluation team was that higher point-of-service quality would lead to higher levels of student engagement (a result that was supported by findings reported in the Interim Report, where PQA scores were found to be related to levels of engagement reported directly by participating youth), and that higher levels of engagement would lead to higher levels of youth participation. Possibly, quality estimates derived from the spring 2011 observations were underpowered in terms of

yielding valid estimates of program quality, given that only four offerings were observed across two days.

Table 11. Summary of Direct Effect of Center and Student-Level Predictors on the Total Hours of Student Participation in ACE Centers, Spring 2011 Sample

Predictors	Coefficient	Std. Err.	p-value
Program Quality Measures			
Observation: High POS Quality Cluster[1]	-31.738	59.186	0.595
Observation: Low POS Quality Cluster[1]	-88.844	56.778	0.127
Observation: High OCE Cluster[1]	-87.268	42.114	0.046*
Staff Survey: Lower Reported Quality Cluster	-90.546	34.729	0.014*
Other Center-Level Variables			
School-Based Center	-116.529	42.598	0.011*
Center Maturity	3.838	26.872	0.888
Staffed Mostly by School-Day Teachers	-23.510	47.693	0.625
Student-Level Predictors			
Grade Level: Middle[2]	26.871	18.387	0.154
Grade Level: High[2]	-200.687	40.372	0.000***
Alignment with High Academic Enrichment Profile	-125.446	162.467	0.446
Alignment with High Homework Help Profile	106.421	126.293	0.406
Activity Category Diversity	22.598	6.0162	0.001**

Note: *statistically significant at 0.05, ** statistically significant at 0.01, *** statistically significant at 0.001
[1] Coefficient values are predicated on a comparison to the mean hours of participation among students enrolled in centers assigned to the *high APT-O/academic climate* cluster.
[2] Coefficient values are predicated on a comparison to the mean hours of participation among students enrolled in elementary students.

In terms of other center-level predictors, only one center characteristic was found to be significantly related to the total number of hours students attended the program. In this regard, students attending centers associated with school-based grantees attended fewer hours of programming than students attending centers associated with non-school-based grantees. This may be due to the fact that in this sample, approximately two thirds of the students served in school-based programs were middle or high school students. In contrast, the programs operated by non-school-based grantees, such as Learning Tree and Big Thought, primarily served elementary students—19% of the students in centers operated by non-school-based grantees were secondary students. Since older students participate in ACE-funded programming less frequently, it is

possible these grade level differences account for the gap between school- and non-school-based programs.

Among the centers visited in spring 2011, only two student-level predictors were found to be significantly related to the total number of hours of program attendance. First, high school students attended fewer hours of programming (an average of 200 hours less) than either middle or elementary school students, a finding that was anticipated. Second, students participating in more categories of activities attended more hours of ACE activities than students involved in fewer types of activities (an average of 23 hours more).

Summary

Many of the analyses described in this section explored the relationship between quality profiles derived from measures implemented with sites visited in the spring of 2011 and the number of hours students at the centers participated in ACE programming in 2010–11. In some instances, quality profiles were found to be related to the levels of program participation in the manner predicted. When compared to a cluster that was moderate quality (the *high APT-O/Academic Climate* cluster), students attending centers in the *lower reported quality* cluster participated in fewer hours of ACE programming than students associated with centers in the *higher reported quality* cluster (an average of 91 hours less). In addition, students in the *high OCE* cluster (indicating high engagement but little substance related to academic skill-building and youth development) had fewer hours of participation in 2010–11 than students enrolled at centers in the moderate quality cluster (an average of 87 hours less).

Duration of Participation in ACE Programming

Several center- and youth-level variables were found to be significantly related to the duration of time students participated in ACE centers in 2010–11 that were visited in spring 2011. (As noted earlier, duration is the number of days between the student's first and last day of participation.) As shown in Table 12, students enrolled in centers assigned to the *high POS quality* cluster attended ACE-funded programming for a significantly longer duration (an average of 82 days longer) than students enrolled in centers assigned to other clusters, and students enrolled in *low POS quality* centers attended programming for a shorter duration (an average of 94 days shorter). These findings indicate a relationship between adoption of the practices described in the PQA, APT-O, and PQA Academic Climate scale and student duration in ACE programming.

This relationship was found to be even stronger for high school students. High school student participation in centers assigned to the *high POS quality* cluster was of longer duration than high school participation in centers assigned to other quality clusters, and high school student participation in *low POS quality* centers was of shorter duration.

(See Appendix D for additional details on cross-level interactions like the one just described).

A series of other center-level predictors were found to be significantly related to the duration of participation in ACE centers.

- *School-Based Centers.* Students attending centers associated with school-based grantees attended programming for a shorter duration than students attending centers associated with non-school-based grantees. As mentioned in the previous section, this marginally significant finding may be related to the fact that a much larger proportion of the students served in school-based programs were either middle or high school students, who, compared to elementary school students, have a demonstrated tendency to participate at lower levels and for a shorter duration. This was particularly the case for students attending school-based centers who were more aligned with the *high homework help* activity profile.

- *Center Maturity.* Students at more mature centers attended ACE programs for a longer duration than students at less mature centers. This was particularly the case for high school students.

- *Mostly Teacher Staffing Model.* Students at centers staffed mostly by school-day teachers attended for a shorter duration than students at centers that were less dependent on school-day teachers. Again, this finding could be related to the fact that school-based centers were more apt than non-school-based centers to rely on teachers to deliver program activities; and school-based centers served a high proportion of secondary students, who tend to participate less frequently and for a shorter duration than elementary students.

Table 12. Summary of Direct Effect of Center and Student-Level Predictors on the Duration of Student Participation in ACE-Funded Programming, Spring 2011 Sample

Predictors	Coefficient	Std. Err.	p-value
Program Quality Measures			
Observation: High POS Quality Cluster[1]	82.217	32.769	0.018*
Observation: Low POS Quality Cluster[1]	-93.828	33.661	0.009**
Observation: High OCE Cluster[1]	26.514	25.724	0.311
Staff Survey: Lower Reported Quality Cluster	-17.679	21.487	0.417

Predictors	Coefficient	Std. Err.	p-value
Other Center-Level Variables			
School-Based Center	-50.889	28.972	0.088[+]
Maturity	72.598	16.935	$p < .001$***
Staffed Mostly by School-Day Teachers	-60.806	26.189	0.027*
Student-Level Predictors			
Grade Level: Middle[2]	-136.724	28.875	0.509
Grade Level: High[2]	-200.687	40.372	$p < .001$***
Alignment with High Academic Enrichment Profile	11.417	75.807	0.882
Alignment with High Homework Help Profile	57.551	78.821	0.471
Activity Category Diversity	29.120	3.320	$p < .001$***

Note: [+]marginally significant at 0.10, *statistically significant at 0.05, ** statistically significant at 0.01, *** statistically significant at 0.001
[1] Coefficient values are predicated on a comparison to the mean duration of participation among students enrolled in centers assigned to the High APT-O/Academic Climate Cluster.
[2] Coefficient values are predicated on a comparison to the mean duration of participation among students enrolled in elementary centers.

Only two student-level predictors were found to be significantly related to the duration of participation by students enrolled in the spring 2011 site visit centers.

- **High School Students.** High school students attended ACE-funded programming for a shorter duration (an average of 136 days less) than either middle or elementary students, a finding that was anticipated.

- **Activity Category Diversity.** Students participating in more categories of activities attended ACE programs for a longer duration (an average of 29 days more) than students involved in fewer types of activities.

The most notable findings related to this set of analyses pertain to the significant relationship between center assignment to the *high POS quality* cluster or *low POS quality* cluster and the duration of student participation. For each cluster, the relationship between program quality and the duration of participation was in the direction hypothesized, with students enrolled in high quality centers demonstrating longer participation than students enrolled in moderate quality centers (assigned to the High APT-O/Academic Climate Cluster). However, the hypothesized negative relationship between centers assigned to the *lower reported quality* cluster and the duration of participation was not evident, nor was a negative relationship between centers assigned to the *high OCE* cluster and participation duration. This latter result is potentially less surprising since it seems reasonable that youth would persist in activities

they find to be highly engaging, even if there is substantively less emphasis on skill-building.

Summary

Two points should be made about the observed relationship between program quality and participation-related outcomes:

- The approach adopted by the evaluation team to construct quality profiles based on observation and staff survey data did not yield a domain of variables that were consistently predictive of both the total number of hours of ACE participation and the duration of participation. While no analyses have been conducted to explore potential reasons for why a lack of consistency was found across the outcomes examined in terms of the predictive value of quality-related profiles, it is suspected that two elements may be at work. The first relates to sources of error that were likely to degrade the robustness of the quality-related measures. In the case of observations, only a small slice of the full domain of programming delivered at a given center during the 2010–11 school year could be observed during the spring 2011 site visits. The absence of data on a broader domain of center offerings across a larger window of service delivery contributed to error in the measures derived from the observations conducted. In terms of the staff survey, the propensity for some respondents to provide responses influenced by social desirability or a misunderstanding of the practices embedded in survey questions also contributed to measurement error by some order of magnitude.

  In addition, it is unclear how parent motivations for having their children enrolled in ACE programming influences participation levels irrespective of program quality, particularly at the elementary level where parents may perceive student enrollment in programming as primarily a school-age child care option. Parent motivation in this regard could serve to weaken the relationship between program quality and levels of participation in ACE-funded activities.

- However, each analysis resulted in significant findings that supported, at some level, the hypothesized relationship between program quality and participation outcomes. The importance and relevance of this finding should not be understated and has ramifications for how TEA should conceptualize and structure its supports designed to support quality in ACE-funded programs.

The Relationship of Center Quality to Student Outcomes

The results described in the previous section indicate a likely relationship between research-based, quality practices and student participation in ACE programs. The

question is, what evidence exists that center implementation of quality practices has a positive impact on student outcomes?

In order to answer to this question, a three-stage analytic strategy was developed to assess how program quality was related to the effect of participating in ACE programs on a variety of student outcomes. The outcomes were those of students in Grades 4 through 12 in sites visited in spring 2011, and were associated with the 2010–11 program year. Each of the outcomes was selected because of its connection to the core goals and objectives of the ACE program. They include: TAKS-Reading/Language Arts and Mathematics scores; the number of school-day absences; the number of disciplinary incidents; and grade level promotion.

In the first stage, a propensity score matching analysis was undertaken. In any evaluation of a program where participants are not randomly assigned to participate or not participate in the program, the problem of selection is paramount. When comparing outcomes for students who participated in the program versus those who did not, the assumption is that students who participated in the ACE program were different from those who did not attend. One difference was the student's (or their parent's) decision to participate in the ACE program. Also, ACE programs targeted certain types of students (primarily at-risk students), and certain types of students were more likely to accept the offer to participate (e.g., students who, for one reason or another, determined that they were likely to benefit from the program). These differences among students can bias estimates of program effectiveness because they make it difficult to disentangle preexisting differences between participating and non-participating students from the effect of attending the program.

Propensity score matching (described in more detail in Appendix E) was used to address this problem. Propensity score stratification is a statistical technique that allows a comparison of outcomes among students who are similar on all available baseline characteristics, including past academic performance. The analysis matched students who participated in the spring 2011 site visit centers, and attended center activities for 30 days or more in 2010–11, with students from the same feeder schools who did not participate in the program. With data on all of the students' characteristics related to their decision to participate in the program and their outcomes, this quasi-experimental design allowed an estimation of the causal effect of participating in the ACE program during the 2010–11 school year.

Thirty days was defined as "treatment" in this set of analyses, ensuring that any comparison of program effect was based on students having significant exposure to the ACE program activities. The 30-day participation threshold has been utilized by the U.S. Department of Education since the 21st CCLC program's inception to define what constitutes regular attendance in the program. The 30-day threshold has also been shown in other statewide evaluations of 21st CCLC to be associated with positive and

significant effects on state assessment scores, particularly in mathematics (Naftzger, Vinson, & Swanlund, 2012; Naftzger, Vinson, Manzeske, & Gibbs, 2011).

Once a sample of non-program participants was identified, analyses were conducted using multilevel modeling techniques to calculate an effect estimate for each *individual center* on each of the outcomes under consideration. Then, the *effect size* of the center on the outcome in question was used as an outcome variable in a multiple regression analysis where each of the following quality variables were predictors:

- Membership in the *high POS quality* cluster based on observation data

- Membership in the *low POS quality* cluster based on observation data

- Membership in the *high OCE* cluster based on observation data

- Membership in the *lower reported quality* cluster based on staff survey data

The goal of the analyses was to understand whether the predictor variables related to program quality actually predicted the effect of centers on student outcomes. (The center-level effect size[28] served as the dependent variable in these analyses). As described in previous sections of this report, variables related to program quality were constructed to support the classification of each center represented in the spring 2011 sample into a quality profile group, ranging from low overall quality to high overall quality. Membership in a given profile type could then be used as a predictor in models constructed to assess if a relationship existed between centers assigned to a given quality profile and student outcomes. This approach allowed the following research question to be answered: "Does the impact on student outcomes vary by relevant ACE program characteristics, including center quality?" It is important to note that while the center-level estimates on student outcomes were *causal*, in that they depicted a given center's impact on the outcomes under consideration, the analyses exploring the relationship between program quality and center-level effect sizes were only correlational in nature. In this sense, causal inferences cannot be drawn about how program quality led to positive program outcomes. Nevertheless, this set of analyses is considered to be unique within the 21st CCLC field, and the findings are potentially useful to both TEA and the afterschool community in general.

---

[28] Effect size refers to the magnitude of the relationship between two variables, in this case program participation and outcome. A small effect size indicates a somewhat weak relationship. In the Interim Report, for example, participation in the program, for students in Grades 4–12, had an effect size of .027 on TAKS-ELA/Reading and of .032 on TAKS-Math. This is a mid-score range for an effect size. To interpret the meaningfulness of such an effect size, it is important to know that generally at the mid score range of the scale score distribution, the difference between one correct score is 7 scale score points; therefore, the effect size for TAKS-Math translates to ACE participants scoring higher than similar but non-participating students by half of one question, or 3.5 scale score points at the mid-score range.

Other center-level variables incorporated into the regression models were selected because of their significance in either the Interim Report or other statewide evaluations of the 21st CCLC program. These were school-based status (whether the center was associated with a school-based grantee); center maturity; and staffing model (whether the center relied primarily on school-day teachers to instruct activities).

Results are presented first for academic outcomes (TAKS assessment scores and grade promotion) and then for behavioral outcomes (disciplinary incidents and school-day attendance).

Academic Outcomes

Three academic outcomes were examined: TAKS-Reading/ELA scores, TAKS-Mathematics scores, and grade promotion. As shown in Table 13, quality-related predictors were found to be significant in the analyses associated with center-level effect sizes related to TAKS-Reading/ELA performance and grade promotion. Centers assigned to the *low POS quality* cluster were found to have significantly lower effect sizes on the TAKS-Reading/ELA scores than centers assigned to other quality clusters. This finding was consistent with what had been hypothesized by the evaluation team.

On the other hand, centers assigned to the *high POS quality* cluster did not have higher effect sizes[29] in terms of TAKS results than other centers. The centers in this cluster did, however, have significantly higher effect sizes in relation to supporting student grade promotion than centers in other clusters. No relationship was found to exist between center membership in a given quality cluster and TAKS-Mathematics performance, nor was membership in the *high OCE* cluster or the *lower reported quality* cluster associated with any of the outcomes examined in Table 13.

In terms of the other center-level predictors, centers associated with school-based grantees were found to have higher effect sizes in terms of TAKS-Reading/ELA and Mathematics achievement than non-school-based centers. (This finding was moderately significant - $p < .10$.) Of some interest was that centers staffed mostly by teachers demonstrated lower effect sizes in terms of supporting grade promotion than centers employing a different type of staffing model. (This finding was also moderately significant.) This finding was not anticipated by the evaluation team.

---

[29] Effect size refers to the magnitude of the relationship between two variables, in this case program participation and outcome. In general, effect sizes in educational research do not exceed 1.0.

Table 13. Summary of Direct Effect of Center-Level Predictors on Academic Outcomes, 2010–11 School Year

Predictors	TAKS Reading			TAKS Mathematics			Grade Promotion		
	Coef	Std. Err.	p-value	Coef	Std. Err.	p-value	Coef	Std. Err.	p-value
Program Quality Measures									
Observation: High POS Quality Cluster[1]	-0.150	0.184	0.422	-0.079	0.127	0.539	0.897	0.364	0.020*
Observation: Low POS Quality Cluster[1]	-0.475	0.209	0.031*	-0.148	0.144	0.313	0.281	0.413	0.501
Observation: High OCE Cluster[1]	0.018	0.168	0.916	0.103	0.116	0.382	0.113	0.332	0.737
Staff Survey: Lower Reported Quality Cluster	0.210	0.130	0.117	0.122	0.090	0.185	0.319	.257	.224
Other Center-Level Variables									
School-Based Center	0.269	0.152	0.088[+]	0.183	0.105	0.093[+]	-0.235	0.301	0.443
Maturity	0.126	0.089	0.166	0.051	0.061	0.416	0.069	0.176	0.697
Staffed Mostly by School-Day Teachers	0.078	0.150	0.607	0.050	0.104	0.631	-0.575	0.297	0.062[+]

Note: [+]marginally significant at 0.10, *statistically significant at 0.05

[1] Coefficient values are predicated on a comparison to the mean effect sizes associated with centers assigned to the High APT-O/Academic Climate Cluster.

Behavioral Outcomes

The behavioral outcomes examined in this analysis are school absences and disciplinary incidents during the school day. As shown in Table 14, quality-related predictors were found to be significant only in analyses associated with effects sizes related to decreases in disciplinary incidents. Centers assigned to the *high POS quality* cluster and the *high OCE* cluster were found to have higher effect sizes related to a decrease in disciplinary incidents than centers assigned to other cluster types. (The coefficients in Table 14 are negative since the desired outcome was a decline in disciplinary incidents.)

A series of moderately significant relationships were also found between several center-level predictors and center effect sizes on the behavioral outcomes examined:

- More mature centers had higher effect sizes in reducing disciplinary incidents than less mature centers. This finding was anticipated by the evaluation team

and was based on the hypothesis that centers become more effective in implementing programs over time.

- The centers staffed mostly by school-day teachers were found to have lower effect sizes in reducing disciplinary incidents than centers with a different staffing approach.

Table 14. Summary of Direct Effect of Center-Level Predictors on Behavioral Outcomes, 2010–11 School Year

Predictors	School Absences			Disciplinary Incidents		
	Coef	Std. Err.	p-value	Coef	Std. Err.	p-value
Program Quality Measures						
Observation: High POS Quality Cluster[1]	-0.075	0.233	0.748	-0.722	0.322	0.033*
Observation: Low POS Quality Cluster[1]	-0.239	0.264	0.372	-0.489	0.365	0.191
Observation: High OCE Cluster[1]	-0.108	0.212	0.613	-0.635	0.293	0.039*
Staff Survey: Lower Reported Quality Cluster	-0.069	0.164	0.675	0.262	0.277	.257
Other Center-Level Variables						
School-Based Center	-0.061	0.192	0.754	0.029	0.266	0.916
Maturity	-0.063	0.112	0.581	-0.269	0.155	0.094[+]
Staffed Mostly by School-Day Teachers	-0.107	0.189	0.577	0.537	0.262	0.050[+]

Note: [+]marginally significant at 0.10, *statistically significant at 0.05
[1] Coefficient values are predicated on a comparison to the mean effect sizes associated with centers assigned to the High APT-O/Academic Climate Cluster.

Summary

One of the unique characteristics of the 2010–11 evaluation of the ACE programs was the availability of multiple measures of program quality on a random sample of centers operating during the 2010–11 school year. In addition, data were available on academic and behavioral outcomes for both ACE program participants and non-participants.

The underlying hypothesis guiding the analyses was that higher levels of program quality would be positively related to participation, academic, and behavioral outcomes. Variables related to quality were developed from observation and staff survey measures—each considering the extent to which staff implemented practices supportive of academic skill-building and mastery, and youth development. Variables indicative of both *high* and *low* program quality were found to be related to the domain of outcomes examined in the manner hypothesized.

Centers assigned to the *high POS quality* cluster were characterized by high program quality on each of the three observation protocols employed during the spring 2011 visits: the PQA; the APT-O/PQA Academic Climate Scale; and the OCE. Outcomes related to center membership in the *high POS quality* cluster included:

- Students enrolled in centers assigned to the *high POS quality* cluster attended ACE programs across a longer span of days than students enrolled in centers assigned to other clusters.

- Centers assigned to the *high POS quality* cluster were found to have higher effect sizes than other centers in terms of supporting a decrease in disciplinary incidents.

- Centers assigned to the *high POS quality* cluster were found to have higher effect sizes than other centers in terms of supporting student grade promotion.

There were two clusters indicating centers were low quality. The first was the *low POS quality* cluster, based on levels of performance on each of the observation instruments used during the spring 2011 visits. The second was a *lower reported quality* cluster, based on data collected from the staff survey where mean scores indicated a lower degree of adoption of quality practices to support academic skill-building and youth development. The hypothesis of the evaluation team was that students enrolled in low quality centers would participate at lower levels, and that the centers would have less of an impact on academic and behavioral outcomes. For the low quality clusters, three outcomes related to center membership in the clusters supported the hypothesis:

- Students attending centers in the *lower reported quality* cluster attended fewer hours of ACE program activities than students associated with centers in the *higher reported quality* cluster.

- Students enrolled in *low POS quality* centers attended activities for a shorter span of time in 2010–11 than students in centers in the other observation-based clusters.

- Centers assigned to the *low POS quality* cluster had lower effect sizes in terms of supporting student performance on the TAKS-Reading/Language Arts assessment than centers assigned to other quality cluster types.

Seven participation, academic, and behavioral outcomes were examined, and only two were found *not* to have a significant relationship with either high or low quality-related predictors as hypothesized: school-day attendance, and TAKS-Mathematics performance. That is not to say that ACE-funded programming did not impact these outcomes in a positive fashion. Results highlighted both in the Interim Report and Chapter 5 of this report indicate that participation in ACE programs positively impacted both school-day attendance and TAKS-Mathematics performance. What was not found

was a relationship between higher levels of program quality as measured by the evaluation team and improved student performance on these outcomes.

Summary of Findings

Ultimately, the results described in this chapter indicate a relationship between higher levels of program quality and outcomes related to student participation and student outcomes in the manner hypothesized. What was not demonstrated was a consistent relationship between a given predictor of program quality and the outcomes examined.

It is important to note that the quality-related analyses undertaken explored the correlational associations between program quality and other center-related characteristics on the one hand, and a variety of participation, academic, and behavioral outcomes on the other. Findings are correlational and descriptive in nature, and causal inferences cannot be drawn about how program quality leads to positive program outcomes. Nevertheless, the results outlined in this section are encouraging for two reasons:

- The observation and staff survey measures adopted during the spring of 2011 led to the formation of quality profiles that in some instances were correlated with participation, academic, and behavioral outcomes as hypothesized. This suggests that further application of these methods and measures in future rounds of evaluation of ACE programs may generate estimates of quality that will be useful in exploring the relationship between program quality and student outcomes.

- Each of the measures used to formulate quality estimates are related to specific practices that program staff can adopt to support quality programming that fosters academic skill-building and youth development. This information may be helpful for TEA as the agency designs and delivers training, professional development, and technical assistance intended to improve program quality.

The next chapter will present a replication of Year 1 impact analyses, but in this case, program participation will be defined as occurring once a student participates for 30 or 60 days. Analyses will be predicated on comparisons between participants and non-participants and between high program attendees and low program attendees. Outcomes to be considered include TAKS Reading/ELA and Mathematics scores, school-day absences, disciplinary incidences, and grade level promotion.

Chapter 5
Program Impact on Student Outcomes

One of the primary objectives of the ACE evaluation is to understand the relationship between participation in ACE programs and student improvement, particularly improvement on outcomes related to academic performance, school-day attendance, disciplinary incidents, and promotion rates. It is these outcomes toward which ACE programs are to direct their programming.

In the analysis that was presented in the Interim Report (April 2012, referenced earlier), a propensity score stratification approach was used to assess the impact of the ACE program on student performance on the spring 2010 TAKS-Reading/ELA and TAKS-Mathematics assessments, as well as on behavioral outcomes associated with the 2009–10 school year (campus discipline incidents and days, and school-day absences). The stratification approach facilitated a comparison of outcomes for ACE program participants with students who were similar in all observable ways except program attendance. Unlike the analyses outlined in Chapter 4, the goal of the analyses described in the Interim Report was to estimate the causal impact of ACE programming on student outcomes. Impact analyses performed in relation to outcomes associated with the 2009–10 school year demonstrated that ACE program participation in 2009–10 was associated with higher TAKS scores in Reading/ELA and Mathematics; fewer assigned disciplinary days and fewer disciplinary incidents (in Grades 9–12 only); and fewer school-day absences. Although each of these findings was statistically significant, the effect sizes were relatively small. Findings associated with the non-academic outcomes (i.e., outcomes other than improvement in TAKS scores) were discussed in terms of either the decreased rate of occurrence for the number of discipline incidences and absences or the increased odds of being promoted to the next grade level.

One of the limitations associated with the analyses undertaken in the Interim Report was that any student who participated in ACE programming for at least one day during the 2009–10 school year was counted as a program participant. Generally, this threshold was too low in terms of how much participation in ACE programming one would expect before a meaningful impact on student outcomes would be achievable. As a result, during the second year of the ACE evaluation, an effort was made to replicate the impact analyses outlined in the Interim Report, but to define ACE participation as having taken place at either 30 or 60 days. It was hypothesized that the impact of the ACE program would be greater for students who attended the program more frequently. This approach allowed the evaluation team to answer the following questions:

- To what extent do students who have higher participation rates demonstrate better academic and behavioral outcomes as compared with similar students who participate in ACE programming at lower levels?

- To what extent do students participating in services and activities funded by 21st CCLC demonstrate better achievement (along with other student outcomes) as compared to similar students not participating in the program?

From the inception of the 21st CCLC initiative, the U.S. Department of Education defined 30 days of participation as regular program attendance. Consistent with this definition, only outcomes of students who participated in the ACE programs for 30 or more days were considered for this report. In addition to the 30-day participation threshold, the analysis presented in this chapter explored the impact of a higher level of program participation on student outcomes. Specifically, the 30-day threshold was doubled to 60 days to assess the impact of higher levels of participation.

Thus, the analysis assessed the relative impact on outcomes of greater levels of participation in the program—30 to 59 days versus 60 days or more—and differences in outcomes between ACE program participants and a matched sample of students from the same schools who did not participate in the program. Variables included in these analyses are limited to the data that were available for each type of student. Analyses comparing outcomes for program participants and non-participants were run separately for students attending ACE programming for 30 and 60 days.

Data Sources and Methods

Propensity score matching (described in more detail in Appendix E) was used to address the problem of selection bias first described in Chapter 4. With data on all of the students' characteristics related to their decision to participate in the program and their outcomes, this quasi-experimental design allowed for an estimation of the causal effect of participating in the ACE program during the 2010–11 school year. This approach was employed to explore differences in outcomes between high program attendees (those attending 60 days or more) and low program attendees (those attending 30 to 59 days), and when assessing the impact of the program on student outcomes as compared to students who did not participate in the ACE program.

Data on student academic and behavioral outcomes were obtained from the TEA-maintained data warehouse, including TAKS and PEIMS. More specifically, the academic and behavioral outcomes examined in this section in relation to students enrolled in all Cycle 5 and 6 centers include:

- TAKS-Reading/ELA and Mathematics scores for the 2010–11 school year
- The number of school-day absences during the 2010–11 school year
- The number of disciplinary incidents during the 2010–11 school year

- Grade level promotion for the start of the 2011–12 school year[30]

In the following sections, findings on the effect of participating in the ACE program on 2010–11 outcomes are presented, in terms of both the effect of high program attendance relative to low program attendance, and the effect of the program relative to non-participants.

Program Impact: Comparing High Attendees with Low Attendees

In this section, findings compare outcomes of high and low attendees: students who attended 60 or more days of ACE programming and similar students who attended 30 to 59 days of programming. This information will be especially useful for TEA and ACE program staff in determining if keeping students involved in ACE programs for 60 or more days results in improved student outcomes.

Separate hierarchical linear regression models were constructed for each grade level to examine the effect of high program attendance on student outcomes.[31] Impact estimates for each grade level were then pooled to create a single, weighted average for each outcome. This approach is described in more detail in Appendix E. Table 15 shows the number of high and low attendee cases analyzed by grade level. As the table shows, in Grades 4–8, there were between 4,651 and 6,936 student cases per grade level in the high attending group. In the high school grades, there were between 1,232 and 1,723 students per grade level participating in the high attending group. Although low attendees exceeded high attendees in the high school grades, the opposite was true in Grades 4–8, where high attendees outnumbered low attendees.

---

[30] Analyzing the effect of ACE program participation on grade promotion was not conducted for the Interim Report; it is new to the current report.
[31] Both TAKS outcomes were analyzed using hierarchical linear regression models. The additional academically-related outcomes were analyzed using hierarchical generalized linear models. That is, the variables for the number of disciplinary incidences and the number of days absent were modeled assuming a Poisson distribution. The variables for whether a student was promoted to the next grade were modeled assuming a Bernoulli distribution.

Table 15. Number of High Attendee (60 days or more) and Low Attendee (30 to 59 days) ACE Student Cases Analyzed, by Grade Level, 2010–11

	High Attendees (60 Days or More)	Low Attendees (30-59 Days)	Total
Grade 4	6,936	2,579	9,515
Grade 5	6,373	2,703	9,076
Grade 6	6,642	4,458	11,100
Grade 7	5,235	4,670	9,905
Grade 8	4,651	4,255	8,906
Grade 9	1,683	2,610	4,293
Grade 10	1,723	2,309	4,032
Grade 11	1,572	2,383	3,955
Grade 12	1,232	2,019	3,251

Source: TX21st

The Effect of High Attendance on TAKS Scores

Analyses examining the impact of high attendance (60 or more days) on TAKS scores demonstrated that high attendance had a positive and significant effect on both TAKS-Reading/ELA and TAKS-Mathematics scores. Students in Grades 4–12 who participated in the program for 60 days or more scored higher on the TAKS assessment outcomes than similar students who participated in the program for 30 to 59 days. Higher levels of attendance in the program for students in Grades 4–12 had an effect size of .022 on TAKS-ELA/Reading and .024 on TAKS-Mathematics (see Table 16). (Effect sizes are presented in standard deviation units.) The magnitude of these effect sizes is very small and, from a practical standpoint, insignificant. To interpret the meaningfulness of such an effect size, it is important to know that at the scale score range where a student moves to the *met standard* level, the difference between one correct score is 12 scale score points;[32] therefore, the effect size for TAKS-Reading/ELA translates to ACE participants scoring higher than similar but non-participating students by nearly one-fifth of one question, or 2.2 scale score points at this range.[33]

---

[32] The score intervals become substantially larger at lower and higher ends of the distribution.

[33] Stated differently, the standard deviation for the Grade 4 TAKS-Reading/ELA assessment is 100 scale score points. Given the effect size of .022, 2.2 percent of the standard deviation is 2.2 sale score points.

Table 16. Effect of High ACE Program Attendance on 2010–11 TAKS Outcomes Relative to Students Participating at Lower Levels

Group	TAKS-ELA/Reading				TAKS-Math			
	Effect Size	Std. Err.	t	p-value	Effect Size	Std. Err.	T	p-value
Grades 4–12	0.022	0.005	3.930	<.001**	0.024	0.006	4.367	<.001**
Grades 4–5	0.025	0.012	2.081	0.038*	0.019	0.012	1.619	0.106
Grades 6–8	0.022	0.008	2.076	0.007**	0.025	0.008	3.041	0.002**
Grade 9[+]	0.033	0.017	2.001	0.046*	0.054	0.018	3.045	0.002**
Grades 9–10[+]	0.039	0.012	3.311	<.001**	0.052	0.012	4.292	<.001**

Source: TEA TAKS data

Note: *statistically significant at 0.05, **statistically significant at 0.01.

Note: [+] Grades 11 and 12 were not included in this set of analyses given that very few students in these grade levels took TAKS during the 2010–11 school year. Grade 9 is presented separately given that most high schools students take TAKS in grade 9.

The Effect of High Attendance on Discipline

Analyses examining the impact of high program attendance on disciplinary incidents demonstrated that higher attendance in the ACE program had a significant effect on reducing the number of discipline incidents (see Table 17) associated with regular school-day attendance. For students in Grades 4–12, participating in the ACE program at a higher rate decreased the rate of disciplinary incidents by 18%. The magnitude of this effect size is small to moderate. Again, these results were fairly consistent across each of the grade level categories examined.

Table 17. Effect of High ACE Program Attendance on School-Day Discipline Incidences Relative to Students Participating at Lower Levels

Group	Number of Discipline Incidences			
	Rate Ratio	Effect Size	Std. Err.	p-value
Grades 4–12	-18%	-0.199	0.011	<.001***
Grades 4–5	-19%	-0.215	0.036	<.001**
Grades 6–8	-18%	-0.203	0.013	<.001**
Grades 9–12	-16%	-0.179	0.023	<.001**

Source: TEA PEIMS data, 2010–11

Note: **statistically significant at 0.001

The Effect of High Attendance on Absences

Analyses examining the impact of high-level program attendance on school-day absences showed that a higher level of attendance in the ACE program had a significant effect on reducing the number of school-day absences (see Table 18). For students in Grades 4–12, high attendance in an ACE program decreased the rate of being absent by 14%. The magnitude of this effect size is small to moderate. Results were fairly consistent across each of the grade level categories examined.

Table 18. Effect of High ACE Program Attendance on School-Day Absences Relative to Students Participating at Lower Levels

Group	Absences			
	Event Ratio	Effect Size	Std. Err.	p-value
Grades 4–12	-14%	-0.151	0.004	<.001**
Grades 4–5	-14%	-0.148	0.008	<.001**
Grades 6–8	-13%	-0.138	0.005	<.001**
Grades 9–12	-16%	-0.170	0.006	<.001**

Source: TEA PEIMS data, 2010–11
Note: **statistically significant at 0.01

The Effect of High Attendance on Grade Promotion

Analyses examining the impact of high-level program attendance on grade promotion demonstrated that higher attendance in the ACE program had a significant effect on increasing the rate of grade promotion. For students in Grades 4–11, participating in the ACE program at a higher rate increased the rate of grade promotion by 30%. As shown in Table 19, for each grade level breakout, the effect of higher attendance on grade promotion was also significant, with the largest effects associated with students in Grades 4–5, where the rate of grade promotion increased by 40% among the high attending group. Nevertheless, the magnitude of the effect sizes outlined in Table 19 are small to moderate, suggesting that grade promotion was one outcome where higher levels of attendance in ACE programming had a significant and potentially meaningful impact.

Table 19. Effect of High ACE Program Attendance on Grade Promotion Relative to Students Participating at Lower Levels

Group	Event Ratio	Grade Promotion		
		Effect Size	Std. Err.	p-value
Grades 4–11[+]	30%	0.263	0.044	<.001**
Grades 4–5	40%	0.334	0.087	<.001**
Grades 6–8	31%	0.268	0.071	<.001**
Grades 9–11	23%	0.210	0.071	.003**

Source: TEA PEIMS data, 2010–11
Note: ***statistically significant at 0.01
Note: [+] Grade 12 was not considered in this analysis since the concept of graduation has a different connotation and a set of criteria that are more detailed and complex than simply moving from one grade level to the next.

Summary of High ACE Attendance Impact Analysis Findings

The results highlighted in this section demonstrate that higher levels of attendance in ACE-funded programs may have increased levels of TAKS-Reading/ELA and Mathematics performance, reduced disciplinary incidents and school-day absences, and supported grade promotion. However, the effect of higher levels of attendance on TAKS-related outcomes, discipline, and absences were quite small. Higher levels of attendance in ACE programs proved to be more impactful in terms of supporting grade promotion. Students attending 60 days or more had a rate of grade promotion 23% to 40% higher than students attending 30-59 days.

Here again, the issue of program quality warrants consideration. As shown in Chapter 4, higher quality programs were more apt to be associated with a greater number of total days of attendance and participation by students across a longer duration of the school year. Further, students with higher attendance levels showed some differences in outcomes. The differences in outcomes were most pronounced in grade promotion. Differences were less pronounced on outcomes related to TAKS, disciplinary incidents, and school-day attendance.

Program Impact: Comparing Program Participants with Non-Participants

A propensity score stratification approach was also used to assess the impact of the ACE program on student outcomes by comparing ACE program participants (included in the above analyses) with students who were similar in all observable ways except program attendance. As in the above analyses, program participation was defined in two separate ways to create a sharper contrast between participants and non-participants. A group of low attendees was identified as having participated in at least 30 days of programming; these students were compared to students who did not

participate in any ACE programming. A group of high attendees was identified as having participated in at least 60 days of programming; they were also compared to students who did not participate in any ACE programming. In contrast, in the Interim Evaluation Report, program participation was defined as at least one day of attendance in ACE-funded programming.

Separate hierarchical linear regression modeling techniques were conducted for *low* (30 to 59 days of attendance) and *high* (at least 60 days of attendance) attendees, by grade level, to examine the effect of the ACE program on non-participating students.[34] Impact estimates for each grade level were then pooled to create a single, weighted average for each outcome. This approach is described in more detail in Appendix E. Table 20 shows the number of low attendees by grade level. As the table shows, in Grades 4–8, there were between 8,905 and 11,065 student cases per grade level in the ACE programs. In the high school grades, there were between 3,237 and 4,289 students per grade level. For some grade levels, the comparison groups comprised of students not participating in the program were approximately five to seven times larger than the treatment groups.

---

[34] Both TAKS-Reading/ELA and TAKS-Mathematics outcomes were analyzed using hierarchical linear regression models. Just as with the within-program analyses, the additional academically-related outcomes were analyzed using hierarchical generalized linear models. That is, the variables for the number of disciplinary incidences, and the number of days absent, were modeled assuming a Poisson distribution. The variable for whether a student was promoted to the next grade was modeled assuming a Bernoulli distribution.

**Table 20. Number of Low Attendee ACE Student Cases Analyzed
by Grade Level, 2010–11**

	Treatment	Comparison	Total
Grade 4	9,510	21,538	31,048
Grade 5	9,073	26,082	35,155
Grade 6	11,065	63,264	74,329
Grade 7	9,901	12,746	22,647
Grade 8	8,905	12,656	21,561
Grade 9	4,289	31,074	35,363
Grade 10	4,026	23,088	27,114
Grade 11	3,924	4,386	8,310
Grade 12	3,237	3,390	6,627

Source: TX21st; TEA PEIMS data, 2010–11

Notes: The treatment counts in the above table are different from the sum of High and Low attendees from Table 15 because they represent analytic samples from two different propensity score analyses. Each propensity score model restricts the range of data to student cases that have overlapping propensity scores; therefore, identical cases may not be analyzed in separate models. Comparison counts are higher than treatment counts because a propensity score stratification approach was used, as opposed to a one-to-one matching approach. The stratification approach uses as many of the comparison cases as possible.

Table 21 shows the number of high attendees included in the analysis by grade level. As the table shows, in Grades 4–8, there were between 4,649 and 6,927 student cases in the ACE programs. In the high school grades, there were between 1,222 and 1,722 students participating. Because a propensity score stratification approach (as opposed to a one-to-one matching approach) was used, whereby as many comparison cases as possible that are similar to the treatment cases are utilized, for some grade levels the comparison groups were significantly larger than the treatment groups (see, for example, Grade 10).

Table 21. Number of High Attendees ACE Student Cases Analyzed by Grade Level, 2010–11

	Treatment	Comparison	Total
Grade 4	6,927	26,644	33,571
Grade 5	6,368	24,996	31,364
Grade 6	6,630	9,267	15,897
Grade 7	5,224	11,557	16,781
Grade 8	4,649	48,213	52,862
Grade 9	1,683	42,814	44,497
Grade 10	1,722	44,196	45,918
Grade 11	1,553	6,603	8,156
Grade 12	1,222	33,730	34,952

Source: TX21st; TEA PEIMS data, 2010–11

For additional descriptive information about treatment and comparison group performance on the outcomes examined in this chapter, please see Appendix F. In the following section, outcomes of ACE program participants are compared with students in a non-participant comparison group.

Program Impact: Comparing Participants with Non-Participants

The second set of analyses presented in this chapter compared outcomes of ACE program participants with a non-participant comparison group. Outcomes were based on 2010–11 data, and included academic outcomes as measured on TAKS scores, discipline, absences, and grade promotion. Findings corresponding to TAKS outcomes are discussed in terms of standard deviation units; findings associated with the other outcomes are discussed in terms of either the decreased rate of occurrence for the number of discipline incidences and absences or the increased odds of being promoted to the next grade level.

Program Effect on TAKS-ELA/Reading Scores

Analyses examining the impact of program attendance compared participants to a non-participant comparison group and found that participation in the ACE program had a positive and significant effect on TAKS-ELA/Reading scores. Both high and low program attendees scored higher on the TAKS assessment outcomes than similar students who did not participate. Low participation in the program (at least 30 days) for students in Grades 4–12 had an effect size of .041 on TAKS-ELA/Reading, whereas high participation (at least 60 days) for student in Grades 4–12 had an effect size of .03 (see Table 22). (Effect sizes are presented in standard deviation units.)

It is important to note that while significant findings were observed for TAKS-ELA/Reading in Grades 4–12, significant findings were not observed in Grades 4–8 for ELA/Reading. Therefore, the significant findings are related to program effects on Grades 9–10, with an effect size of .192 and .195 for low and high attendees, respectively. These results are shown in Table 22.

Similar results were found in the Interim Report in relation to TAKS-Reading/ELA outcomes: program effects were not significant for students in Grades 4–8 but were for Grades 9–10, although the program effect was significantly smaller at .063 when treatment was defined as at least one day of ACE participation. These results suggest the relative benefit of higher levels of participation in ACE-funded programming on TAKS-ELA/Reading performance for high school students participating for 30 days or more.

Table 22. Effect of ACE Program Participation on 2010–11 TAKS-ELA/Reading for Low and High Attendees, Relative to Non-Participants

Group	TAKS-ELA/Reading – Low Attendees				TAKS-ELA/Reading – High Attendees			
	Effect Size	Std. Err.	t	p-value	Effect Size	Std. Err.	T	p-value
Grades 4–12	0.041	0.005	8.780	<.001**	0.030	0.006	5.101	<.001**
Grades 4–5	-0.011	0.010	-1.047	0.295	0.002	0.011	0.136	0.892
Grades 6–8	-0.004	0.007	-0.539	0.590	0.011	0.009	1.317	0.188
Grades 9–10[+]	0.192	0.010	18.523	<.001**	0.195	0.020	9.967	<.001**

Source: TEA TAKS data, 2010–11
Note: **statistically significant at 0.01
Note: [+] Grades 11 and 12 were not included in this set of analyses given that very few students in these grade levels took TAKS during the 2010–11 school year.

Program Effect on TAKS-Mathematics Scores

Analyses examining the impact of program attendance also compared program participants (both high and low attendees) to a non-participant comparison group, and found that participating in the ACE program had a positive and significant effect on TAKS-Mathematics scores. Both low- and high-attending students who participated in the program scored higher on the TAKS assessment outcomes than similar students who did not participate. Low participation in the program (at least 30 days) for students in Grades 4–12 had an effect size of .041 on TAKS-Mathematics, whereas high participation (at least 60 days) for student in Grades 4–12 had an effect size of .029 (see Table 23). (Effect sizes are presented in standard deviation units.) The magnitude of these effect sizes is very small and, from a practical standpoint, insignificant.

Just as with the TAKS-Reading/ELA finding, it is important to note that while significant findings were observed for TAKS-Mathematics in Grades 4–12, significant findings were not observed in Grades 4–8 for Mathematics. Therefore, the significant findings are related to program effects on Grades 9–10, with an effect size of .187 and .186 for low and high attendees, respectively. These results are shown in Table 23. The magnitude of these effect sizes is moderate.

Similar results were found in the Interim Report in relation to TAKS Mathematics outcomes, where program effects were largest for students in Grades 9–10, although the program effect was significantly smaller at .057 when treatment was defined as at least one day of ACE participation. Here again, these results suggest the relative benefit of higher levels of participation in ACE-funded programming on TAKS-Math performance for high school students participating for 30 days or more.

Table 23. Effect of ACE Program Participation on 2009–10 TAKS Outcomes for High Attendees and Low Attendees, Relative to Non-Participants

Group	TAKS-Math – Low Attendees				TAKS-Math – High Attendees			
	Effect Size	Std. Err.	T	p-value	Effect Size	Std. Err.	t	p-value
Grades 4–12	0.041	0.005	8.706	<.001**	0.029	0.006	5.042	<.001**
Grades 4–5	-0.007	0.010	-0.687	0.492	-0.001	0.012	-0.057	0.954
Grades 6–8	-0.005	0.007	-0.788	0.430	0.008	0.009	0.892	0.372
Grade 9–10[+]	0.187	0.010	17.951	<.001**	0.186	0.019	9.768	<.001**

Source: TEA TAKS data, 2010–11
Note: **statistically significant at 0.01
Note: [+] Grades 11 and 12 were not included in this set of analyses given that very few students in these grade levels took TAKS during the 2010–11 school year.

Program Effect on Discipline

Analyses examining the impact of program attendance compared program attendees to a non-participant comparison group and found that participating in the ACE program had a significant effect on reducing the number of discipline incidents associated with regular school-day attendance (see Table 24). For low and high attendees in Grades 4–12, participating in the ACE program decreased the rate of being assigned disciplinary incidents by 6% and 11%, respectively. However, a significant decrease in disciplinary incidents was only observed in Grades 6–12. For low and high attendees in Grades 6–8, participation decreased the rate of being assigned disciplinary incidents by 3% and 11%, respectively. In comparison, for low and high attendees in Grades 9–12, participation decreased the rate of being assigned disciplinary incidents by 13% and 16%, respectively. The magnitude of the effect sizes for Grades 9–12 is small to moderate. The reduction in the rate of disciplinary incidents outlined in Table 23 was

significantly lower than what was recorded for participants in Grades 9–12 in the Interim Report, where the rate of reduction was only 5%. In addition, while no significant impact was found on reducing the rate of disciplinary incidents for students in Grades 6–8 in the Interim Evaluation Report, the rate of reduction was found to be 11% for students participating for 60 days or more. Here again, these results suggest the relative benefit of higher levels of participation in ACE-funded programming on reducing disciplinary incidents for high school students participating for 30 days or more, and for middle school students participating for 60 days or more.

Table 24. Effect of ACE Program Participation on 2010–11 School-Day Discipline Incidents Relative to Non-Participants

Group	Number of Discipline Incidences – Low Attendees				Number of Discipline Incidences – High Attendees			
	Rate Ratio	Effect Size	Std. Err.	p-value	Rate Ratio	Effect Size	Std. Err.	p-value
Grades 4–12	-6%	-0.058	0.007	<.001**	-11%	-0.119	0.009	<.001**
Grades 4–5	-1%	0.014	0.021	0.511	-4%	-0.040	0.025	0.112
Grades 6–8	-3%	-0.034	0.008	<.001**	-11%	-0.115	0.012	<.001**
Grades 9–12	-13%	-0.142	0.013	<.001**	-16%	-0.174	0.019	<.001**

Source: TEA PEIMS data, 2010–11
Note: **statistically significant at 0.01

Program Effect on Absences

Analyses examining the impact of program attendance compared program participants to a non-participant comparison group and found that participating in the ACE program had a significant effect on reducing the number of school-day absences (see Table 25). For low- and high-attending students in Grades 4–11,[35] participating in the ACE program decreased the rate of being absent by 14% and 15%, respectively. For low and high attendees in Grades 4–5, participating in the ACE program decreased the rate of being absent by 13% and 10%, respectively. For low- and high-attending students in Grades 6–8, participation decreased the rate of being absent by 13% and 15%, respectively. Finally, for low- and high-attending students in Grades 9–11, participation decreased the rate of being absent by 19% and 18%, respectively. The magnitude of the effect sizes for Grades 6–8 and 9–11 is small to moderate. While a similar pattern was found in the Interim Report, the magnitude of the effect sizes for low- and high-attending students were found to be significantly higher for each grade level. Here

---

[35] Due to model specification, the Grade 12 model for low-attending students did not converge and findings are not available; therefore, to compare findings from the same grade levels, Grade 12 estimates are not reported.

again, these results suggest the relative benefit of higher levels of participation in ACE-funded programming on reducing the rate of being absent, particularly for students participating 60-days or more.

Table 25. Effect of ACE Program Participation on 2010–11 School Day Absences

Group	Number of School Day Absences – Low Attendees				Number of School Day Absences – High Attendees			
	Rate Ratio	Effect Size	Std. Err.	p-value	Rate Ratio	Effect Size	Std. Err.	p-value
Grades 4–11	-14%	-0.146	0.002	<.001**	-15%	-0.159	0.003	<.001**
Grades 4–5	-13%	-0.076	0.047	<.001**	-10%	-0.110	0.006	<.001**
Grades 6–8	-13%	-0.136	0.222	<.001**	-15%	-0.165	0.005	<.001**
Grades 9–11[+]	-19%	-0.205	0.521	<.001**	-18%	-0.204	0.010	<.001**

Source: TEA PEIMS data, 2010–11
Notes: **statistically significant at 0.01.
Note: [+] Grades 12 is not included in the table because analyses oriented as assessing program impact on school absences failed to converge in relation to this grade level.

Program Effect on Grade Promotion

The analysis found that participating in the ACE program had a significant effect on increasing the likelihood of being promoted to the next grade level (see Table 26). For low- and high-attending students in Grades 4–5 and 7–11,[36] participating in the ACE program increased the likelihood of being promoted to the next grade level by 43% and 47%, respectively. For high attendees in Grades 4–5, participating in the ACE program increased the likelihood by 18%. For low and high attendees in Grades 7–8, participation increased the likelihood by 26% and 29%, respectively. The magnitude of this effect size is small to moderate.

Although the above effect sizes are not large, the program effect on grade promotion among high school students was much larger. Among students in Grades 9–11, for both low and high attendees, participation in an ACE program increased the likelihood of being promoted to the next grade level by 79% and 97%, respectively. The magnitude of this effect size is large.

However, while these results are encouraging, it is not clear if there are other key student characteristics or attributes that lead high school students to participate in ACE-funded programming and that are less represented in the comparison group. For

---

[36] Due to model specification, the Grade 6 model for high-attending students did not converge and findings are not available; therefore, to compare findings from the same grade levels, Grade 6 estimates are not reported.

example, student motivation is an attribute that could not be taken into consideration when completing the matching process given the unavailability of these data. It could be that high school students participating in ACE programming were simply more motivated to succeed academically. In this regard, while these large effect sizes are encouraging, it should be noted that it could be the case that other student characteristics and attributes are driving this outcome.

Table 26. Effect of ACE Program Participation on 2011–12 Grade Promotion

Group	Grade Promotion – Low Attendees				Grade Promotion – High Attendees			
	Odds Ratio	Effect Size	Std. Err.	p-value	Odds Ratio	Effect Size	Std. Err.	p-value
Grades 4–5, 7–11	+43%	0.354	0.029	<.001**	+47%	0.384	0.040	<.001**
Grades 4–5	+3%	0.028	0.057	0.619	+18%	0.169	0.068	0.014*
Grades 7–8	+26%	0.229	0.057	<.001**	+29%	0.253	0.075	0.001**
Grades 9–11	+79%	0.584	0.041	<.001**	+97%	0.675	0.065	<.001**

Source: TEA PEIMS data, 2011–12
Notes: *statistically significant at 0.05; **statistically significant at 0.01.

Summary

The purpose of this chapter was to explore the relationship between participation in ACE programs and student improvement, particularly improvement on outcomes related to academic performance, school-day attendance, disciplinary incidents, and promotion rates. One set of analyses examined the difference in program impact by comparing high attendees (60 days or more) with low attendees (30 to 59 days). Higher levels of attendance in ACE-funded programs were found to have small to moderate effects on TAKS-Reading/ELA and Mathematics performance, reduced disciplinary incidents and school-day absences, and grade promotion. In particular, higher levels of attendance in ACE programs proved to be most impactful in terms of supporting grade promotion. Students attending for 60 days or more had a rate of grade promotion 23% to 40% higher than students attending 30-59 days.

The second set of analyses outlined in this chapter examined program impact by comparing outcomes for ACE participants with students who did not participate in the program. In this case, impact analyses were done two ways. In one set of analyses, ACE participation was defined as being 30 days or more (low attending). In the second set of analyses, ACE participation was defined as being 60 days or more (high attending). For both low- and high-attending students, ACE program participation had a statistically significant impact on TAKS scores, discipline, absences, and grade promotion for many of the grade levels relative to students who did not participate in the program.

- For Grades 9–12 only, ACE program participation was associated with higher TAKS scores in ELA/Reading and Mathematics.

- For Grades 6–12 only, ACE program participants had fewer disciplinary incidents than non-participating students.

- For low-attending students in Grades 4–5 and high-attending students in Grades 4–11, program participation was associated with fewer school-day absences.

- Finally, program participation was associated with increased likelihood of grade promotion in Grades 6–11 for low-attending students and in Grades 4–11 for high-attending students.

- For both low- and high-attending students, impacts on grade promotion were especially substantial in Grades 9–11.

A key Critical Success Factor articulated in the CSM adopted by TEA for the ACE program outlines the importance of school connectedness and involvement with supporting student academic growth and development. Having seen significant program effects in reducing disciplinary referrals and school-day absences, and particularly in supporting grade promotion in secondary programs, it seems appropriate to examine in greater detail how ACE programming is supporting school connectedness and bonding in a way that supports the achievement of these outcomes. In addition, program effects related to TAKS Reading/ELA and Mathematics for students in Grades 9–10 were noteworthy.

A topic that warrants additional consideration in the Year 3 evaluation of the program is the question of what attracts these older youth to ACE programming and keeps them participating for an extended period of time. This information could be especially helpful to TEA and the field in understanding how to reach and engage older youth in a manner that is likely to lead to positive behavioral and academic outcomes.

Chapter 6
Summary of Findings and Recommendations

There were several characteristics associated with the 2010–12 evaluation of the ACE program that made it unique among statewide evaluations that have been conducted of 21st CCLC programs to date:

- The availability of multiple measures of program quality and the capacity to explore on-the-ground program implementation in great detail;

- Access to data on a variety of academic and behavioral outcomes for both ACE program participants and non-participants;

- Sufficient resources to conduct a series of rigorous analyses examining the relationship between program quality and the achievement of desired student outcomes.

Program quality can be conceptualized as the binding thread that is woven throughout the CSM, adopted by TEA to guide implementation of ACE programming in a way that is likely to impact the lives of participating youth in meaningful and lasting ways. The 2010–12 evaluation was guided by seven research questions, most of which attempt to document what constitutes quality practice in ACE programs and the influence of program quality on student outcomes. Findings related to each of the research questions are presented below.

1. What instructional approaches are associated with high levels of student engagement at the point of service?

The analysis of observation data from both the 2011 and the spring 2012 site visits showed that three instructional approaches distinguished high quality activities. One feature was clarity of purpose, whereby the activities were clearly designed to achieve explicit objectives. In the high quality academic enrichment sessions, the instructors clearly stated the learning objectives and then led students through a variety of learning activities related to those objectives. In the non-academic enrichment sessions, the objectives may not have been as explicitly stated, but the activities were still purposed toward learning within the context of the sessions.

A second feature was intentional use of time. This feature, which was anchored in planning and pacing, was found to be essential for keeping students busy and engaged throughout the observed sessions. Materials were ready when the sessions began. Routines were worked into sessions so little time was wasted when students began sessions and transitioned from one activity to another. The pace was generally quick, and, as a result, student accomplishments by the end of the activities were evident.

The third feature was an active and interactive instructor, who continually engaged with students, even when students were working in small groups or on their own. The instructors moved about the room, looked over students' shoulders, asked questions that deepened student knowledge, noticed and helped when students had trouble, and managed student behaviors before any became disruptive.

2. *What organizational processes are found to be drivers of instructional/point of service quality at high performing centers?*

Several organizational processes were examined to determine whether they were drivers of instructional and point of service quality. Center intentionality was a key dimension that was examined. Center intentionality refers to the center's purpose and the degree to which that purpose is defined and appropriate resources provided. Among the 15 ACE programs in the sample, those with high center intentionality with respect to academic enrichment programming showed a strong focus on meeting academic objectives that were emphasized during the school-day core classes. This was particularly evident in the elementary centers and one middle school center that had developed curricula for their academic enrichment activities and consistently aligned the curricula to school-day learning objectives. Intentionality with respect to non-academic enrichment programming was uneven across all grade levels. Across the centers, there did not appear to be a shared understanding of what non-academic enrichment programming is. Numerous respondents referred to these activities as opportunities to relax, play games, and be physically active—thus, they were non-academic but did not provide opportunities to learn new knowledge and skills associated with a non-academic subject area or discipline.

Practices to monitor for improvement (that is, to improve the quality of activities) were evident in the majority of the centers in the site visit sample. Particularly strong were the centers that systematically approached monitoring of academic enrichment activities and modifying curricula and instructional approaches to better engage students and meet learning objectives. To an extent, this was more obvious among the elementary centers in the sample. In the secondary centers in the sample, where monitoring was assessed as high, attendance was typically the foremost factor in monitoring and making improvements in programming. Site coordinators and, in some cases, project directors considered student attendance and student motivation to participate—that is, were students interested in the activities? Were the sessions themselves appealing? Were there barriers to attendance that needed to be addressed? Responses to monitoring of this type included implementing different types of activities, particularly non-academic enrichment activities, based on student interest, and working with instructors to better align instructional methods with students' developmental needs. This type of monitoring seems appropriate in high schools where students have a high level of choice regarding participation.

Nearly all of the 15 centers showed a clear linkage to the school day. Close ties with the school were reflected in information transfers related to students and school learning objectives, availability of space, and administrative support.

Staff development and staff collaboration were other means of supporting program quality. Among the 15 centers in the 2012 site visit sample, staff development and opportunities for collaboration, particularly collaboration in scheduled sessions, were more available to instructors who were not certified teachers than to those who were teachers during the school day. The school-day teachers had little time to meet formally, although they had some opportunities for sharing information on an informal basis. The scheduling constraints, a site coordinator said, made it difficult to improve instructional methods of staff in order to make them more engaging for youth. In this group of ACE programs, the elementary centers were more reliant on non-certified teachers (usually affiliated with the grantee organization) to instruct sessions than secondary centers, where activities were frequently instructed by certified teachers. Thus, it was elementary centers in the sample—more than secondary centers—that provided opportunities for both staff development and collaboration for at least a majority of instructors. Providing opportunities for staff development and collaboration for certified teachers may be a persistent challenge for many afterschool programs. Sometimes, this is addressed through the professional development of site coordinators, which was extensive for the majority of the centers in the sample.

Finally, community connections were important to many of the centers, most of which had developed partnerships with organizations and agencies that resulted in expanded opportunities for programming, youth, and parents.

3. What innovative strategies and approaches can be identified from these centers that warrant replication and emulation?

Many innovative strategies and approaches from centers serving youth in elementary, middle, and high schools were presented in the Year Two Evaluation Report. They are too numerous to list in this summary chapter, though may be referred to as models that might be replicated. It is important to note that although the type of activities that were observed and presented as exemplars are varied, they all exemplified the three instructional features described in Chapter 2: clarity of purpose, intentional use of time, and an active and interactive instructor.

4. What is the relationship between the characteristics of individual youth, center quality, and other center characteristics and levels of student participation in ACE programming?

Getting students to participate in ACE-funded programming consistently and on a sustained basis over time is a critical first step in enhancing the likelihood that students will achieve desired program outcomes. A hypothesis was that students enrolled in

centers demonstrating higher quality would be more likely to participate in more total hours of programming, and for a longer duration, during the course of the 2010–11 school year (measured by the number of days between the first day of participation and the last day of participation).

In order to test this hypothesis, centers visited in 2011 were classified into different quality profile types based on observation and staff survey data collected during this period. Four profile types were defined based on observation data: 1) *high POS quality*; 2) *low POS quality*; 3) *high APT-O/Academic Climate*; and 4) *high OCE*. It was expected that participation-related outcomes would be better in centers classified in the *high POS quality* cluster, because centers assigned to this cluster were characterized by high program quality on each of the three observation protocols employed: the PQA, the OCE, and the APT-O. In addition, centers assigned to the *low POS quality* cluster were expected to do less well on participation-related outcomes because of lower levels of observed program quality, based on the same three protocols.

When multilevel models were run to explore these relationships, the hypothesized pattern of results was found: students enrolled in *high POS quality* centers participated in programming for a significantly longer duration during the 2010–11 school year, while students enrolled in *low POS quality* centers participated for a significantly shorter duration.

However, a similar result was not found when the total hours of ACE programming attendance was used as an outcome. Neither *high POS quality* nor *low POS quality* was predictive of this outcome.

Surveys completed by activity leaders asked respondents to report the extent to which they engaged in quality-related practices. From the analysis of these responses, two quality-related clusters were identified: (1) *lower reported quality* centers, and (2) *higher reported quality* centers. The hypothesis was that centers in the *higher reported quality* cluster would demonstrate better participation-related outcomes than centers enrolled in the *lower reported quality* cluster. When the total hours of participation in ACE programs was used as an outcome, the hypothesized relationships was found, suggesting that the implementation of higher quality practices was related to more hours of participation. However, implementation of these same practices was not found to be related to the duration of participation.

Overall, the approach adopted by the evaluation team to construct quality profiles based on observation and staff survey data did not yield variables that were consistently predictive of both the total number of hours of ACE participation and the duration of participation. However, each analysis resulted in significant findings that supported, at some level, the hypothesized relationship between program quality and participation outcomes.

5. Does the impact on student outcomes vary by relevant ACE program characteristics, including center quality?

One of the primary objectives of the ACE evaluation is to understand the relationship between participation in ACE programs and student improvement, particularly improvement on outcomes related to academic performance, school-day attendance, disciplinary incidents, and promotion rates. It is these outcomes toward which ACE programs are to direct their programming.

Using the observation and staff survey-based quality clusters employed to answer research question 4, a three-stage analytic strategy was developed to assess how program quality was related to the effect of participating in ACE programming on a variety of student outcomes associated with the 2010–11 programming period:

- TAKS-Reading/ELA and TAKS-Mathematics scores

- The number of school-day absences

- The number of disciplinary incidents

- Grade level promotion

It was hypothesized that centers demonstrating higher quality would have a stronger, positive effect on each of these outcomes, while lower quality centers would have a weaker effect. This hypothesis was borne out in the following findings:

- Centers assigned to the *high POS quality* cluster were found to have *higher* effect sizes in terms of supporting a decrease in disciplinary incidents than centers assigned to other quality types

- Centers assigned to the *high POS quality* cluster were found to have *higher* effect sizes in terms of supporting student grade promotion than centers assigned to other quality types

- Centers assigned to the *low POS quality* cluster were found to have *lower* effect sizes in terms of supporting student performance on the TAKS-Reading/ELA assessment than centers assigned to other quality types.

No relationship was found between higher quality programming and larger effect sizes in terms of program impact on TAKS-Mathematics scores and school-day absences.

6. To what extent do students who have higher participation rates demonstrate better academic and behavioral outcomes as compared with similar students who participate in 21st CCLC at lower levels?

Analyses were undertaken to explore the extent to which students who attended programming for 60 days or more demonstrated better outcomes than similar students who participated in ACE programming for 30 to 59 days. Results from these analyses

demonstrated that higher levels of attendance in ACE-funded programs were associated with higher levels of TAKS-Reading/ELA and Mathematics performance, reduced disciplinary incidents and school-day absences, and grade promotion. However, the effect of higher levels of attendance on TAKS-related outcomes was quite small. Higher levels of attendance in ACE programs proved to be more impactful in terms of reducing disciplinary incidents and school-day absences and particularly in supporting grade promotion. In the latter case, students attending 60 days or more had a rate of grade promotion 23% to 40% higher than students attending 30-59 days. This information provides ACE programs with additional understanding regarding how much additional program impact can be derived from keeping students engaged in ACE programming for 60 days or more.

7. *To what extent do students participating in services and activities funded by 21st CCLC demonstrate better achievement (along with other student outcomes) as compared to similar students not participating in the program?*

Analyses were undertaken to assess the impact of the ACE program on student outcomes by comparing ACE program participants with students who were similar in all observable ways except program attendance. Program participation was defined in two separate ways to create a sharper contrast between participants and non-participants. A group of "low" program participants was identified as having participated in at least 30 days of programming, and these students were compared to students who did not participate in any ACE programming. A group of "high" program participants was identified as having participated in at least 60 days of programming, and these students were also compared to students who did not participate in any ACE programming.

For both low- and high-attending students, ACE program participation had a statistically significant impact on TAKS scores, discipline, absences, and grade promotion for many of the grade levels relative to students who did not participate in the program.

- For Grades 9–12 only, ACE program participation was associated with higher TAKS scores in Reading/ELA and Mathematics.

- For Grades 6–12 only, ACE program participants had fewer disciplinary incidents than non-participating students.

- For low-attending students in Grades 4–5 and high-attending students in Grades 4–11, program participation was associated with fewer school-day absences.

- Program participation was associated with increased likelihood of grade promotion in Grades 6–11 for low-attending students and in Grades 4–11 for high-attending students.

The magnitude of each of these program effects was primarily in the small to moderate range, with the largest effects associated with reductions in school-day absences and

grade promotion. However, for both low- and high-attending students, impacts on grade promotion were especially substantial in Grades 9–11. In this case, participation in an ACE program increased the likelihood of being promoted to the next grade level by 79% and 97%, respectively. The magnitude of this effect size is large.

In addition, while the impact of the program on each of these outcomes was found to be significant in the Interim Evaluation Report, the level of impact was significantly larger for students in both the low- and high-attending groups, demonstrating the importance of retaining students in programming beyond the 30-day threshold.

Primary Themes and Recommendations

Most of the findings outlined in this report can be distilled into two primary themes, both of which can guide future evaluation work and inform how TEA approaches the design and delivery of training, technical assistance, and professional development for staff working in ACE-funded programs:

- *Program Quality Matters.* In some instances, measures of program quality employed during the evaluation to assess center functioning in the adoption of practices to support academic skill-building and youth development were related to both student participation in ACE programs and the achievement of student outcomes. In particular, measures of program quality were found to be related to levels and duration of participation in ACE programming, a decrease in disciplinary incidents, grade promotion, and performance on TAKS-Reading/ELA. Each of the measures employed to formulate quality estimates detail specific practices that program staff can adopt to support implementation of quality programming that supports academic skill-building and mastery, and youth development, among participating students. TEA should consider reviewing these measures to see how the practices articulated in each measure may further inform the formulation and delivery of training, professional development, and technical assistance oriented at supporting centers in improving the quality of their offerings. Each of the tools operationalize the features of high quality activities noted in this report, helping to ensure clarity of purpose and intentional use of time, and to provide markers for the types of behaviors that define an active and interactive instructor.

  In addition, states are increasingly working on the development and implementation of quality assessment tools and mechanisms, like leading indicator systems, to feed data on program quality back to 21st CCLC programs to support quality improvement efforts. Most of these systems are predicated on supporting program adoption of specific quality-related practices. TEA is encouraged to review their current efforts in this regard to see what additional approaches could be implemented to get actionable quality data into the hands of program administrators and staff.

- *High School Students Especially Benefitted from ACE Participation.* Consistently, across each of the outcomes examined, program effects were found to be the greatest for high school students participating in ACE programming. (The 15 centers in the 2012 site visit sample showed that academic support was mainly available through tutorials that were highly aligned to the school day, which may be one reason for the benefits for high school students.) In some instances, these effects were quite large, particularly in relation to grade promotion, where students attending 60 days or more demonstrated a 97% better chance of being promoted to the next grade level as opposed to similar students not enrolled in the program.

  Such results warrant further examination into what attracts these older youth to ACE programming, and what keeps them participating for an extended period of time. Engaging in such an examination is of particular importance since efforts to identify the features of high quality offerings detailed in this report consistently demonstrated lower performance on the part of high school programs on key facets of program quality. This may suggest that there is something unique about high school students who opt to participate in ACE programming that makes it particularly likely they will benefit from their participation in such programming. It is the hypothesis of the research team that student motivation plays a strong role in shaping how high school students interact with and benefit from their participation in ACE programming. Understanding the role motivation plays in how high school students connect with afterschool programming would be especially helpful to TEA and the field in terms of understanding how to reach out and engage older youth in a manner that is likely to lead to positive academic achievement outcomes.

The information in this report provides concrete evidence for, and examples of, how program quality can support the achievement of desired ACE program outcomes. What has not been measured or assessed to date is how youth change as a *direct consequence* of ACE program participation, and how these changes transfer outside the program to impact the types of student academic and behavioral outcomes examined in this report. These more immediate, *within-program outcomes* can fall within a wide spectrum of categories, including social emotional learning; critical thinking and decision-making; initiative and self-direction, and so on (Wilson-Ahlstrom, Yohalem, DuBois, & Ji, 2011). In addition, acquisition of content-specific skills in areas like reading and mathematics are likely to be more targeted in nature within a given ACE program, both in terms of the area of emphasis within a given program and a student's unique area of need, particularly in relation to students falling below proficiency. As we move into Year 3 of the evaluation, steps will be taken to focus more time and attention on understanding how programming impacts more immediate student skills and functioning that translate into desirable academic and behavioral outcomes.

More specifically, with the onset of the 2012–13 school year, TEA will be providing current ACE grantees that have demonstrated a capacity to provide higher quality programming with the opportunity to obtain a State of Texas Assessments of Academic Readiness (STAAR) Supplemental Academic Support Grant to better identify and serve students who are particularly in need of help and support in developing STAAR-related skills. In many respects, the decision by TEA to develop such a program provides a number of opportunities to structure the Year 3 21st CCLC evaluation in a way that allows for a variety of research questions to be addressed that further build on and extend evaluation efforts conducted to date, including:

- How are recipients of STAAR Supplemental Academic Support Grants using these funds to identify and recruit high need students into programming; what steps are being taken to align programming with the CSM adopted by TEA for the 21st CCLC program; and to what extent is programming being delivered in a manner that is consistent with afterschool quality frameworks?

- What characteristics are associated with Supplemental Academic Support Grant-supported activities where there are high levels of youth-reported engagement?

- What impact does programming funded by the STAAR Supplemental Academic Support Grants have on short-term program outcomes, like student task persistence, motivation, and academic self-efficacy?

Taking steps to answer these questions will provide additional valuable information on how the quality of programming funded by ACE leads to important changes in the knowledge and skills of youth that ultimately translate into academic achievement and success.

References

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

Durlak, J. A., & Weissberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills.* Chicago: Collaborative for Academic, Social, and Emotional Learning. Retrieved January 15, 2011, from http://www.casel.org/downloads/ASP-Full.pdf

Ho, D., Imai, K., King, G., & Stuart, E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15,* 199–236. Retrieved January 15, 2011, from http://gking.harvard.edu/files/abs/matchp-abs.shtml

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.

Granger, R. (2008). After-school programs and academics: Implications for policy, practice, and research. *Social Policy Report, 22*(2), 3–19. Ann Arbor, MI: Society for Research in Child Development. Retrieved from http://www.srcd.org/documents/publications/spr/spr22-2.pdf

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

Kane, T. (2004). *The impact of after-school programs: Interpreting the results of four recent evaluations.* Retrieved January 15, 2011, from http://www.wtgrantfoundation.org /publications_and_reports/browse_reports/kane_working_paper

Kim, S.C., & Wilson, M. (2009). A comparative analysis of the ratings in performance assessment using generalizability theory and the many-facet Rasch model. *Journal of Applied Measurement, 10*(4), 408–423.

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

Little, P. (2007). *The quality of school-age care in afterschool settings.* New York and Ann Arbor, MI: National Center for Children in Poverty, Columbia University; Institute for Social Research, University of Michigan.

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

Linacre, J. M. (2009) *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.

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

Naftzger, N., Margolin, J., Kaufman, S., & Ali, A. (2006). *An overview of the 21st CCLC program: 2004–05.* Naperville, IL: Learning Point Associates.

Naftzger, N., Vinson, D., Manzeske, D., and Gibbs, C. (2011). *New Jersey 21st Century Community Learning Centers (21st CCLC) impact report 2009–10.* Naperville, IL: American Institutes for Research.

Naftzger, N., Vinson, M., & Swanlund, A. (2011). *21st Century Community Learning Centers (21st CCLC) analytic support for evaluation and program monitoring: A summary of findings from the collection of student-level data during the 2008-09 reporting period.* Washington, DC: U.S. Department of Education.

Naftzger, N., Vinson, M., & Swanlund, A. (2012). *An evaluation of the Washington 21st Century Community Learning Centers (21st CCLC) program: 2010–11.* Naperville, IL: American Institutes for Research.

Rosenbaum, P.R., & Rubin, D.B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association, 79(387),* 516–524.

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.

Wilson-Ahlstrom, A., Yohalem, N., DuBois, D., & Ji, P. (2011). *From soft skills to hard data: Measuring youth program outcomes.* Washington, DC: Forum for Youth Investment, Impact Strategies.

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

Zief, S. G., Lauver, S., & Maynard, R. A. (2006). *The impacts of after-school programs on student outcomes: A systematic review for the Campbell Collaboration.* Retrieved January 5, 2011, from http://www.campbellcollaboration.org/lib/download/58/

Appendices

Appendix A
Moments of High Engagement: High vs. Low Quality Centers

The following table reflects only moments where engagement was rated as "All" on the Observation of Child Engagement (OCE). The table lists the full domain of activities observed in each category that generated an "All" rating on the OCE. As expected, there were found to be more of these moments associated with activities observed in centers receiving a high quality designation as opposed to centers receiving a low quality designation based on total PQA scores.

Table A1. Comparison of Practices During Moments of "All" Engagement for High and Low Quality Centers

High Total Quality Score and High Engagement Score (OCE = All)	Low Total Quality Score and High Engagement (OCE = All)
Academic Enrichment: High School/Middle School; 1. Students leave personal belongings in hall, put on lab coats upon entry. 2. Staff questioning students, introducing relevant vocabulary in group discussion. 3. Staff reviews steps with groups, groups set about tasks. 4. Students working in small groups, completing tasks, teacher circulates, checking in with student groups. 5. Students enter, collect materials. 6. Students prepare materials for use, teacher circulating, helping when needed. 7. Students compare results. 8. Students reflect on activity. 9. Students watch high interest video (academically themed).	Academic Enrichment: High School/Middle School; 1. Students complete task (light detector) and use. 2. Students enter class and collect materials for task. 3. Students read aloud. 4. Students watch high-interest video (academically themed).

High Total Quality Score and High Engagement Score (OCE = All)	Low Total Quality Score and High Engagement (OCE = All)
Academic Enrichment: Elementary: 1. Teacher begins class with social conversation, then enthusiastically announces it is time for session to begin. 2. Teacher questions for prior knowledge 3. Teacher provides examples of finished product. 4. Teacher checks for understanding 5. Students collect materials from teacher. Teacher makes students wait to begin until all have materials. 6. Teacher uses relevant academic vocabulary to question students about projects. 7. Teacher asks students to describe and reflect on their process. 8. Students compare results.	**Academic Enrichment: Elementary:** 1. Students do activity with sweet treats, during which they answer questions about a story they read. 2. Students play math game where teams work on problems for points. 3. Teacher circulates and checks work, questioning students about their process of finding answers.
Non-Academic Enrichment: **High School/Middle School** 1. Teacher reviews prior knowledge and announces the schedule of the class (warm, informative). 2. Teacher provides tools/materials used by professionals in the field of study for students to use. 3. In a role play, students have choice with respect to companion actors and role play content. 4. Students plan for school event; staff member is the note taker and provides supplemental information regarding logistics only. 5. Staff works as an equal in student planning. Students do not embrace staff suggestion, staff does not use position to insist. 6. Democratic decision-making is evident. 7. Staff direct students in a play. Staff model role and express enthusiasm. — 8. Staff questions students for prior knowledge. 9. Staff demonstrates strong knowledge in subject area. 10. Staff discusses motivation for student actions (in context of play).	**Non-Academic Enrichment:** **High School/Middle School** 1. Students work on craft project. 2. Students engaged in social conversation. 3. Students finish craft projects, clean up and go outside. 4. Students play various board games or work on computers individually or in pairs. 5. Students work at computers, wearing headphones—teacher circulating, looking at screens. 6. Teacher talking with individual students.

High Total Quality Score and High Engagement Score (OCE = All)	Low Total Quality Score and High Engagement (OCE = All)
Non-Academic Enrichment: Elementary 1. Class starts on time, teacher welcomes students. 2. Teacher checks for prior knowledge of session activity/content. 3. Teacher questions students, using student answers as means to deeper questions, more complex aspects of topic: teacher able to deeply discuss content in a variety of ways that employs student prior knowledge. 4. Teacher allows student choice within context of topic (choose how students will approach activity; process choices). 5. Teacher gives specific praise for student work. 6. Teacher connects student activity work to larger community (kids can enter school-wide contest with the product of the day's activity). 7. Teacher organizes clean-up, students work together to clean-up and finish activity/session together. 8. The room is organized into stations, choice of activity. 9. Variety of activities are available—allow for different student strengths (multiple intelligences). 10. Despite high noise level, all students engaged in paired or group activities. 11. Teacher help students try different activities—ensure fairness by making sure no students dominate any one activity, move students at regular intervals. 12. Staff participates in activities with students.	Non-Academic Enrichment: Elementary 1. Session begins on time. 2. Supplies set up for students to begin activity. 3. Staff play holiday-themed music in background (complements activity/craft content). 4. Staff have social conversation with students . 5. Activity is holiday-themed eatable craft. 6. Staff uses student names. 7. Staff circulates among students. 8. Staff gives specific praise. 9. Staff begins students in circle, then organizes participants into groups. 10. Staff gets student attention with clapping (call & response style). 11. Staff gives whole group a drawing activity—one student draws for 10 seconds (measured with a stop watch) then passes on to another student to continue.

High Total Quality Score and High Engagement Score (OCE = All)	Low Total Quality Score and High Engagement (OCE = All)
Homework Help: High School/Middle School 1. Staff circulates as students work. 2. Staff invites students lingering in hallway into session by asking if they have anything they would like to work on. 3. Staff allows students to leave class to turn in an assignment to a teacher (staff trusts/allows students to leave). 4. Staff allow independent work, but oversee students – when one student gives a good definition of a word, staff praises, even though staff was not being spoken to: sense that staff is attentive without intervention. 5. Staff are smiling and available. 6. Staff allow a variety of activities—varying noise levels. 7. Staff support a variety of activities. 8. Staff intervenes when students appear to get off track by entering conversation with specific work-related questions/prompts. 9. Staff tells student not to copy each other's work and explains why they should not do this. 10. Teacher finds a problem that several students are having trouble with and helps them work through it. 11. Teacher prompts with questions, does not do problem for them. 12. Teacher sees several students are reading a book for a class, initiates a discussion of the book. 13. Teachers constantly moving and interacting with students. 14. Teachers aware of students' work, progress, affect. 15. Teachers find ways to be involved in students' work, not satisfied to allow students to work in silence. 16. Teachers help students deepen engagement with their work. 17. Teachers establish environment of active involvement with schoolwork. 18. Staff help students use entire time, don't allow students to pack up early.	Homework Help: High School/Middle School ***There were no moments of "All" engagement as rated by the OCE in the HS/MS "Low" group**

High Total Quality Score and High Engagement Score (OCE = All)	Low Total Quality Score and High Engagement (OCE = All)
Homework Help: Elementary	Homework Help: Elementary
1. Teacher picks up students in class and brings them to computer lab for HH session.	***There were no moments of "All" engagement as rated by the OCE in the elementary "Low" group**
2. Teacher addresses students by name.	
3. Teacher informs students of schedule for session, links content to their school-day class.	
4. Teacher organizes group into class – teaches concept to whole groups for first half of session.	
5. Teacher works a problem on board with whole group contributing.	
6. Teacher employs student helpers, instructs them to "Guide kindly."	
7. Helper students are circulating and helping while the rest of students stay at computer terminals, working on problem.	
8. Teacher provides fun, but academically oriented activity (word puzzle) for students who have finished.	
9. Teacher maintains a quiet environment, tells student helper "you can't help someone if you can't be calm."	
10. Teacher addresses group to correct a mistake that she notices on several student papers	
11. Teacher checks progress: each student comes to desk to speak privately with teacher for 5 minutes.	
12. Students are encouraged to go to the bathroom when they need to.	
13. Teacher checks student progress, allows those who are finished to use computers in pairs or in groups.	
14. Staff treats students warmly.	
15. Staff provides support materials.	
16. Students work in pairs on assignment.	
17. Staff keeps students aware of time: "you have 30 minutes to finish."	
18. Staff provides as much help as students need.	
19. Staff and student clean up a mess together.	

Appendix B. Measures Used to Create Quality Profiles

As noted in the main narrative, both observation and staff survey data collected during 2011 were used to create quality profiles.

Observations

Three observation instruments guided each observation: the Youth Program Quality Assessment /School-Age Program Quality Assessment (PQA), portions of the Assessment of Afterschool Practices Observation Tool (APT-O) that address academic content, and the Observation of Child Engagement (OCE).

The Youth PQA is a validated instrument for observing program activities that serve youth in Grades 4–12, and the School-Age PQA is used to observe activities that serve youth in Grades K–6. The tools were developed by the High/Scope Educational Research Foundation and currently are supported by the Weikart Center, a partner on this project. Both versions of the PQA measure afterschool programming at the point of service, where youth and program staff intersect for instruction and learning. Constructs represented on the tool pertain to how supportive, interactive, and engaging the activity is for participating youth; the extent to which desired pedagogical methods are demonstrated by staff; and the extent to which developmentally appropriate opportunities are afforded to participating youth. The Youth PQA served as the foundation of a large-scale afterschool program improvement intervention tested in four states with funding from the William T. Grant Foundation; the intervention was found to significantly improve the quality of afterschool activities on the constructs measured by the Youth PQA (Smith et al., in review). It is currently the official quality assessment tool for 21st CCLC for several states. Some Texas grantees use the PQA as a self-assessment tool to guide program improvement efforts.[37]

Academic content was also examined through Academic Climate scale of the PQA, which was employed to assess the level of challenge; opportunities to use higher-order thinking skills; and staff feedback, questioning, and guidance of discussion in a way that supported student learning.

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 CCLCs. The APT-O was selected to supplement the PQA because it includes measures of skill-building in reading and mathematics. Although the APT-O

---

[37] See http://etools.highscope.org/pdf/YouthPQA.pdf for information regarding the Youth PQA. See http://www.arbetterbeginnings.com/downloads/SAToolkit/EA7.pdf for sample items from the School-Age PQA.

addresses a number of afterschool quality constructs, only those scales related reading and mathematics skill-building were used for the spring 2011 observations.

The degree to which children are engaged in observed activities was assessed by trained raters using a modified version of the OCE (Rimm-Kaufman & Pianta, 2005), an adaptation of the NICHD Early Child Care Research Network Classroom Observation Scale. The measure consists of five items, *engagement, attention, self-reliance, compliance,* and *disruptive behavior.* Each was rated on a four-point Likert-type scale. The OCE was selected because it provides a behavioral measure of student engagement that is more context-specific and emergent than the other two instruments. The OCE was used to identify episodes during the observed activities in which students were actively engaged in program activities, and then document corresponding instructional practices, grouping strategies, and interactions, and learning opportunities. Observers rated student engagement, and also completed an observation narrative, in which they recorded instructor and student activities, and quotations which illustrated the interactions among students and between the students and instructor.

Staff Survey

The purpose of the online staff survey was to obtain information from frontline staff in the 40 site visit centers who work directly with youth. A particular focus of the survey was on staff use of quality practices that support both positive academic and youth development outcomes. Scales on the survey included were meant to assess how well staff were implementing quality practices related to academic skill-building and youth development in the following areas:

- Intentionality in activity and session design
- Practices supportive of academic skill-building, including linkages to the school day and using data on student academic achievement to inform programming
- Practices supportive of positive youth development
- Opportunities for youth ownership

Appendix C. Rasch Models: Survey and Observational Data

The purpose of this appendix is to supply additional information about the Rasch models employed in the Year Two Evaluation Report. 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 less skilled persons, Rasch modeling techniques take person and item difficulty estimates yielded from an instrument, transform them using a log function, and display them on a logit scale that allows person and item difficulties to be directly compared.

One of the benefits from the application of 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 effectively employed in supporting the domain of analyses needed for the year one report, Rasch analysis techniques were employed to create scale scores for scales associated with several instruments used to support data collection efforts in the spring and fall of 2011. 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 survey and took the following form:

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

    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 (MFRM) (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. Observation of Child Engagement (OCE)

The MFRM 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

In terms of reliability, the Rasch rating scale model allows for the production of indices that indicate the replicability of the model-based respondent ability estimates across similar instruments (Bond & Fox, 2007). As Bond and Fox note, person reliability is enhanced if there is relatively small error in the ability estimates associated with respondents, which in turn is impacted by the number of items used to support the analysis. In Tables C1 and C2, the reliability indices for each individual subscales calibrated across each of the measures employed in the year two report are outlined

and ranged from .66 to .91. Reliability estimates in the range of .60 to .70 are considered minimally acceptable (Bond & Fox, 2007).

In addition, Tables C1 and C2 also include a column labeled *Mean Standardized Outfit*. Values found in this field are useful in assessing the extent to which the data associated with a given analysis fit the Rasch model. Using information about how an individual respondent scored across the full domain of items on a given instrument and how the full sample of respondents scored on a given item, the Rasch model constructs an expected score for each person on each item represented in a given analysis. This expected score is then compared with the observed score for that person on the item in question and a residual is calculated.

These values represent the standardized value of the mean squared residual among the items represented in a given analysis and serves as an indication of model fit. If the data perfectly conformed to the Rasch model, then the mean standardized outfit would be 0. Negative values indicate that there was less variation in the data than expected while positive values indicate more variation was found in the data than expected (Bond & Fox, 2007). The standardized mean outfit values outlined in Tables C1 and C2 are almost all negative, which may suggest that there is redundancy among the items and that a more parsimonious presentation of items on some instruments may be warranted in the future. According to Linacre (2009), non-standardized mean outfit values between .5 and 1.5 are indicative of a productive measure and deemed to be an indicator of acceptable fit, which is the case for each of the subscales outlined Tables B1 and B2.

Table C1. Person Reliability Indices and Outfit Values by Subscale for Scales Calibrated with the Rasch Rating Scale or Dichotomous Model

Instrument/Scale	Rasch Person Reliability Index	Standardized Mean Outfit (Non-standardized)
Staff Survey		
Intentionality in activity and session design	.80	-0.3 (1.01)
Practices supportive of academic skill-building	.84	-.01 (1.04)
Practices supportive of positive youth development	.81	-0.1 (1.02)
Opportunities for youth ownership	.81	-0.3 (.95)

Table C1. Person Reliability Indices and Outfit Values by Subscale for Scales Calibrated with the Rasch Rating Scale or Dichotomous Model (continued)

Instrument/Scale	Rasch Person Reliability Index	Standardized Mean Outfit (Non-standardized)
Observation Tools		
PQA Academic Climate	.70	0.0 (.87)
APT-O – Academic Content	.70	0.0 (.93)

Table C2. Person Reliability Indices and Outfit Values by Subscale for Scales Calibrated with Many Facet Rasch Measurement

Instrument/Scale	Rasch Person Reliability Index	Standardized Mean Outfit (Non-standardized)
PQA		
Total	.91	-0.1 (1.03)
Supportive Environment	.79	-.01 (.96)
Interaction	.69	-0.3 (.85)
Engagement	.66	0.0 (.97)
OCE	.85	0.0 (1.12)

Also, it is important to note that MFRM techniques can also be employed to identify and quantify various sources of error variation. MFRM accomplishes this task by employing fit statistics and separation reliability indices akin to those described earlier to estimate parameters for a specific facet independent of the other facets included in the model. For example, the basic Rasch model allows for both (a) the estimation of the ability of an individual respondent and (b) the difficulty of an individual item and the production of individual standard errors for both persons and items. MFRM allows a researcher to add additional facets to the Rasch model, like rater for example, resulting in the estimation of individual rater severity estimates and standard errors on the same logit scale as person ability and item difficulty estimates, allowing for direct comparison across the three facets in question.

As noted by Kim and Wilson (2009), this feature of MFRM allows the researcher to assess the impact of error variance within each facet on the respondent's ability estimate. In this sense, the probability that a respondent will receive a given score on the measure of interest is a function of the difference between the person's ability and

the difficulty of the task, after adjusting for error introduced by a given measurement facet (like rater severity, for example). In this regard, as Kim and Wilson emphasize, what MFRM yields is an estimate of the respondent's score that is as free as possible from the particularities associated with the measurement facets included in the model.

Capitalizing on this characteristic of MFRM, a partially-crossed method was employed when conducting PQA-related observations in the spring of year one where observers were paired in an intentional manner to allow for the PQA measures to be calibrated using MFRM. In doing so, the evaluation team was able to obtain an estimate of whether a given rater was systematically more lenient or severe in their ratings and adjust calibrated scores to account for this systematic bias demonstrated by the rater. The same raters were employed in conducting fall 2011 visits, preserving the partially-crossed nature of the spring and fall datasets even though only one site visitor was sent to a given site during the fall visits.

In Table C3, the severity measure for each rater involved in the collection of PQA data is outlined. A value of 0 would indicate a completely unbiased rater while negative values indicate a more lenient rater and positive values indicate a more severe rater. Most raters were within .5 logits of 0, so a wide range in severity was not considered to be an issue with these data. In Table C3, outfit values are also presented. Values in this column can be interpreted in the same fashion as those values appearing in Table C1 and C2. Here, however, three raters were found to have both nonstandardized and standardized outfit values beyond desirable levels (raters 1, 4, and 7), which suggests they were not using the PQA rating scale in a consistent fashion across observations. This is different than the issue of systematic bias, which can be quantified and accounted for in the calibration process. Outfit is indicative of error that is being introduced by inconsistent use of the rating scale, although overall reliability levels were still within acceptable ranges as outlined in Table C2. This error cannot be corrected through MFRM; however, this is important information for the evaluation team to have in hand as we prepare to retrain and certify observers, providing us with information about which raters we especially need to target in these efforts.

Table C3. MFRM PQA Severity Measures and Outfit Values

Rater	Severity Measure	Standardized Mean Outfit (Non-standardized)
Rater 1	-.59	4.3 (1.49)
Rater 2	-.25	1.2 (1.05)
Rater 3	-.07	-3.8 (.85)
Rater 4	.04	6.7 (1.89)
Rater 5	.13	2.2 (1.30)
Rater 6	.32	-2.7 (.89)
Rater 7	.42	3.7 (1.55)

In addition to adjusting respondent scores to account for measurement error created by a given facet, MFRM also produces a series of statistics summarizing the outcome of what is termed a fixed effect hypothesis test (Linacre, 2009) for each facet included in the model and their interactions. These tests, employing a chi-square-based statistic, assess whether or not the elements associated with the facet in question can be considered as sharing the same measure after allowing for measurement error. For example, these fixed hypothesis tests allow questions like *Can these raters be thought of as equally lenient?* to be answered. In the case of the PQA data, the answer to this question was *no*, the raters in question could not be considered to be equally lenient ($p < .001$, Chi-Square = 271.3, df = 6). Failure to employ MFRM techniques would have resulted in biased estimates of activity functioning.

Similar techniques were employed with OCE data, although the facet added to the model was not rater, but segment. The OCE instrument was scored every 10 minutes for a given activity, resulting in up to 6 scored segments per observation. MFRM was used to determine if certain segments of an activity were systematically resulting in higher scores, which was found to be the case ($p < .001$, Chi-Square = 40.1, df = 5). Segment four received the highest OCE scores, while segments five and six at the end of the activity received the lowest. This information is important to have in hand when the number of segments that can be observed varies from one activity to another.

Appendix D. Hierarchical Linear Models Employed in Program Quality Analyses

The purpose of this appendix is to further outline the domain of hierarchical linear models run to explore the relationship between center and student characteristics, particularly program quality, and student ACE program participation outcomes in relation to those centers subjected to site visits in the spring of 2011. These analyses were oriented toward addressing the following question: What is the relationship between the characteristics of individual youth, center quality, and other center characteristics and levels of student participation in ACE programming?

Two types of models were run to address this question:

1. *Total Number of Hours of ACE Programming as the Outcome.* The total number of hours of ACE programming (HOURS) served as the outcome of interest at level one in the model outlined below, with level one predictors including the student's alignment with the high enrichment (HIENRICH) or high homework help (HIHOMEWO) profile; the number of activity categories the student was enrolled in during the 2010–11 school year (DIVERSIT); and whether or not the student was either a middle (MIDDLE) or high school (HIGH) student. Level two predictors included center membership in each of the observation-based program quality cluster (HIPOSQUA, LOPOSQUA, and LOSUBQUA); center membership in the staff survey-based low quality cluster (LOSTAFFQ); the school-based status of the grantee associated with the center (SCHOOL_B); how mature the center was in terms of years of operation (MATURITY); and whether or not the center employed a mostly teachers staffing model (MOSTLY_T). This model took the following form:

Level 1 Model

$$\text{HOURS} = \beta_0 + \beta_1(\text{HIENRICH}) + \beta_2(\text{HIHOMEWO}) + \beta_3(\text{DIVERSIT}) + \beta_4(\text{MIDDLE}) + \beta_5(\text{HIGH}) + r$$

Level 2 Model

$$\beta_0 = Y_{00} + Y_{01}(\text{HIPOSQUA}) + Y_{02}(\text{LOPOSQUA}) + Y_{03}(\text{LOSUBPOS}) + Y_{04}(\text{LOSTAFFQ}) + Y_{05}(\text{SCHOOL\_B}) + Y_{06}(\text{MATURITY}) + Y_{07}(\text{MOSTLY\_T}) + u_0$$

$$\beta_1 = Y_{10} + Y_{11}(\text{HIPOSQUA}) + Y_{12}(\text{LOPOSQUA}) + Y_{13}(\text{LOSUBPOS}) + Y_{14}(\text{LOSTAFFQ}) + Y_{15}(\text{SCHOOL\_B}) + Y_{16}(\text{MATURITY}) + Y_{17}(\text{MOSTLY\_T}) + u_1$$

$$\beta_2 = Y_{20} + Y_{21}(\text{HIPOSQUA}) + Y_{22}(\text{LOPOSQUA}) + Y_{23}(\text{LOSUBPOS}) + Y_{24}(\text{LOSTAFFQ}) + Y_{25}(\text{SCHOOL\_B}) + Y_{26}(\text{MATURITY}) + Y_{27}(\text{MOSTLY\_T}) + u_2$$

$$\beta_3 = Y_{30} + Y_{31}(\text{HIPOSQUA}) + Y_{32}(\text{LOPOSQUA}) + Y_{33}(\text{LOSUBPOS}) + Y_{34}(\text{LOSTAFFQ}) + Y_{35}(\text{SCHOOL\_B}) + Y_{36}(\text{MATURITY}) + Y_{37}(\text{MOSTLY\_T}) + u_3$$

$$\beta_4 = Y_{40} + Y_{41}(HIPOSQUA) + Y_{42}(LOPOSQUA) + Y_{43}(LOSUBPOS) + Y_{44}(LOSTAFFQ) +$$
$$Y_{45}(SCHOOL\_B) + Y_{46}(MATURITY) + Y_{47}(MOSTLY\_T) + u_4$$

$$\beta_5 = Y_{50} + Y_{51}(HIPOSQUA) + Y_{52}(LOPOSQUA) + Y_{53}(LOSUBPOS) + Y_{54}(LOSTAFFQ) +$$
$$Y_{55}(SCHOOL\_B) + Y_{56}(MATURITY) + Y_{57}(MOSTLY\_T) + u_5$$

2. *Participation Duration as the Outcome.* The number of calendar days between the student's first day of participation in the ACE programming during the 2010–11 school year and their last day of participation served as the outcome of interest at level one in the model outlined below (FIRSTDAY), while the domain of level one and level two predictors were the same as those listed in relation to model 1.

Level 1 Model

$$FIRSTDAY = \beta_0 + \beta_1(HIENRICH) + \beta_2(HIHOMEWO) + \beta_3(DIVERSIT) + \beta_4(MIDDLE) + \beta_5(HIGH) + r$$

Level 2 Model

$$\beta_0 = Y_{00} + Y_{01}(HIPOSQUA) + Y_{02}(LOPOSQUA) + Y_{03}(LOSUBPOS) + Y_{04}(LOSTAFFQ) +$$
$$Y_{05}(SCHOOL\_B) + Y_{06}(MATURITY) + Y_{07}(MOSTLY\_T) + u_0$$

$$\beta_1 = Y_{10} + Y_{11}(HIPOSQUA) + Y_{12}(LOPOSQUA) + Y_{13}(LOSUBPOS) + Y_{14}(LOSTAFFQ) +$$
$$Y_{15}(SCHOOL\_B) + Y_{16}(MATURITY) + Y_{17}(MOSTLY\_T) + u_1$$

$$\beta_2 = Y_{20} + Y_{21}(HIPOSQUA) + Y_{22}(LOPOSQUA) + Y_{23}(LOSUBPOS) + Y_{24}(LOSTAFFQ) +$$
$$Y_{25}(SCHOOL\_B) + Y_{26}(MATURITY) + Y_{27}(MOSTLY\_T) + u_2$$

$$\beta_3 = Y_{30} + Y_{31}(HIPOSQUA) + Y_{32}(LOPOSQUA) + Y_{33}(LOSUBPOS) + Y_{34}(LOSTAFFQ) +$$
$$Y_{35}(SCHOOL\_B) + Y_{36}(MATURITY) + Y_{37}(MOSTLY\_T) + u_3$$

$$\beta_4 = Y_{40} + Y_{41}(HIPOSQUA) + Y_{42}(LOPOSQUA) + Y_{43}(LOSUBPOS) + Y_{44}(LOSTAFFQ) +$$
$$Y_{45}(SCHOOL\_B) + Y_{46}(MATURITY) + Y_{47}(MOSTLY\_T) + u_4$$

$$\beta_5 = Y_{50} + Y_{51}(HIPOSQUA) + Y_{52}(LOPOSQUA) + Y_{53}(LOSUBPOS) + Y_{54}(LOSTAFFQ) +$$
$$Y_{55}(SCHOOL\_B) + Y_{56}(MATURITY) + Y_{57}(MOSTLY\_T) + u_5$$

Results from each multilevel model are shown in Table D1. It is important to note that only significant cross-level interactions from each model have been displayed.

Table D1. Summary of Results of Center and Student-Level Predictors on Participation-Related Outcomes

Predictors	Hours of Student Participation			Duration of Student Participation		
	Coefficient	Std. Err.	p-value	Coefficient	Std. Err.	p-value
Program Quality Measures						
Observation: High POS Quality Cluster[1]	-31.738	59.186	0.595	82.217	32.769	0.018*
Observation: Low POS Quality Cluster[1]	-88.844	56.778	0.127	-93.828	33.661	0.009**
Observation: High OCE Cluster[1]	-87.268	42.114	0.046*	26.514	25.724	0.311
Staff Survey: Lower Reported Quality Cluster	-90.546	34.729	0.014*	-17.679	21.487	0.417
Other Center-Level Variables						
School-Based Center	-116.529	42.598	0.011*	-50.889	28.972	0.088[+]
Center Maturity	3.838	26.872	0.888	72.598	16.935	0.000***
Staffed Mostly by School-Day Teachers	-23.510	47.693	0.625	-60.806	26.189	0.027*
Student-Level Predictors						
Grade Level: Middle[2]	26.871	18.387	0.154	10.200	15.278	0.509
Grade Level: High[2]	-200.687	40.372	0.000***	-136.724	28.875	0.000***
Alignment with High Academic Enrichment Profile	-125.446	162.467	0.446	11.417	75.807	0.882
Alignment with High Homework Help Profile	106.421	126.293	0.406	57.551	78.821	0.471
Activity Category Diversity	22.598	6.0162	0.001**	29.120	3.320	0.000***

Note: [+]marginally significant at .10; *statistically significant at 0.05, ** statistically significant at 0.01, *** statistically significant at 0.001
[1] Coefficient values are predicated on a comparison to the mean hours of participation among students enrolled in centers assigned to the *high APT-O/academic climate* cluster.
[2] Coefficient values are predicated on a comparison to the mean hours of participation among students enrolled in elementary students.

Table D1. Summary of Results of Center and Student-Level Predictors on Participation-Related Outcomes (Continued)

Predictors	Hours of Student Participation			Duration of Student Participation		
	Coefficient	Std. Err.	p-value	Coefficient	Std. Err.	p-value
Cross-Level Interactions[3]						
Alignment with High Homework Help Profile x Observation: Low POS Quality Cluster[1]	971.172	461.477	0.043*	472.489	293.227	0.117
Grade Level: Middle[2] x Center Maturity	-54.890	25.605	0.040*	-7.841	18.391	0.672
Grade Level: Middle[2] x Staffed Mostly by School-Day Teachers	-100.040	46.298	0.038*	-36.308	33.583	0.288
Grade Level: High[2] x School-Based Center	272.810	103.543	0.013*	50.756	89.291	0.573
Grade Level: High[2] x Center Maturity	221.858	73.714	0.006**	328.544	53.761	$p < .001$***
Alignment with High Homework Help Profile x School-Based Center	-609.236	363.616	0.103	-436.810	231.281	0.068+
Activity Category Diversity x School-Based Center	19.253	16.046	0.239	17.835	9.060	0.057+
Activity Category Diversity x Staffed Mostly by School-Day Teachers	23.993	15.050	0.120	18.349	8.293	0.034*
Grade Level: High[2] x Observation: High POS Quality Cluster[1]	198.135	122.400	0.115	307.250	91.520	0.002**
Grade Level: High[2] x Observation: Low POS Quality Cluster[1]	-10.506	137.018	0.940	-232.026	108.661	0.040*
Grade Level: High[2] x Staffed Mostly by School-Day Teachers	-117.023	108.616	0.290	196.717	80.238	0.020*

[3]Only significant cross-level interactions are shown.

Appendix E. Propensity Score Matching Methods

Propensity score matching is a two-stage process. In the first stage, the probability that each student participates in the ACE program (or in the case of the high versus low attendee analysis, participates in ACE programming for 60-days or more) is modeled on available observable characteristics. By modeling selection into the program, this approach allows us to compare participating and non-participating students (or high versus low attendees) who would have a similar propensity to select into the program based on observables. In the second stage, the predicted probability of participation is used to model student outcomes.

Stage 1: Creation of the Comparison Group. The outcome of interest in modeling propensity scores is treatment status (1 for students participating in the ACE program, 0 for the comparison group or 1 for higher attendees, 0 for low attendees). To account for this binary outcome, logistic regression is used to model the logit (or log-odds) of student group assignment status. Because characteristics of students and the campuses they attend will influence whether they attend the ACE program or participate at higher levels, data on all of these pre-treatment characteristics were acquired from TEA. Student level variables that were used to fit the propensity score models include, but are not limited to, the following:

- Age
- Gender
- Race/ethnicity
- Special education
- Limited English proficient (LEP) status
- Gifted education status
- Previous retention
- Number of prior year disciplinary incidences
- Number of days absent during the prior year
- TAKS scores from the three previous years
- Economically disadvantaged

Campus characteristics used to fit the single-level propensity score model included, but were not limited to the following:

- Attendance rate
- Class size
- Teacher education
- Student mobility

- Percent race/ethnicity
- Percent LEP
- Percent special education
- Accountability status
- Number of full time teachers
- Teacher's average years of experience
- Teacher/student ratio
- Percent economically disadvantaged
- Percent bilingual
- Number of students

Data were not available for each of these covariates for all students. To account for this, indicator variables were used to model the relationship between the pattern of missing data and propensity to participate in the summer program (Rosenbaum & Rubin, 1984).

All pre-treatment covariates were initially considered as candidates for inclusion in the propensity score model. To select an initial propensity score model, we began by regressing each of the covariates on ACE program participation. All covariates with a p value of less than 0.05 were then included in a forward stepwise regression function to produce an initial propensity score model. This approach was used to limit collinearity and include only those variables that were related to program participation.) Propensity scores and propensity score logits were then estimated using this model. We examined overlap in the treatment and comparison groups and deleted non-overlapping cases. We then looked at balance across the two groups on all covariates. Balance statistics (standardized mean differences and variance ratios) were used to guide model selection. The final models included between 34 and 71 covariates, and the adjusted standardized mean differences between the treatment and comparison groups were below 0.2 on all pretreatment covariates, consistent with current best practice in the propensity score literature (Ho, Imai, King, & Stuart, 2007).

Stage 2: Statistical Modeling of Student Outcomes. Outcomes of students in the ACE program were then compared with the outcomes of students who did not participate (the comparison group) or between high (60 days or more) and low (30 to 59 days) attendees. We balanced pretreatment group differences in observed covariates 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 pre-treatment measure of the outcome were also 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

$$y_{ij} = \beta_{0j} + \beta_{1j}21CCLC\ Participation_{ij} + \sum_{S=2}^{15} \beta_s L_{sij} + \beta_{16j}LP_{ij} + \beta_{17j}Pretest_{ij} + r_{ij}$$

Where y_{ij} are the student level outcomes (TAKS scores, discipline, attendance rates, and grade promotion), $21CCLC\ Participation_{ij}$ is an indicator of whether the student participated in the 21st CCLC program (or was a high attender), L_{sij} is an indicator variable for the logit propensity score stratum, LP_{ij} is the logit propensity score, and $Pretest_{ij}$ is the pre-treatment 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 above Level 2 equation includes only β_{0j} because the above hierarchical linear model is a random intercept model; all other coefficients (i.e., participation indicator, logit propensity score stratum, logit propensity score, and pre-treatment 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 above, it is not necessary to include these variables in the final outcome model.

Weighted Averages of Impact Estimates

Analyses were run separately by grade and then pooled together to develop overall estimates of program effect. TAKS results were standardized before pooling to account for scale differences between grades (effect sizes and standard errors were divided by within grade standard deviation). To calculate pooled estimates, the following weighted average equations were used:

Weights for each grade-level were calculated by using the inverse variance (1 divided by the squared standard error of the effect). The following equation shows how a weight is calculated for each grade level g. The weights are calculated such that the sum of the w_g across all grades equals 1.

$$w_g = \frac{\sigma_g^{-2}}{\sum_g \sigma_g^{-2}}$$

In the above equation, σ_g^{-2} is the inverse variance association with the effect for grade g. Using these weights, the pooled effect δ_p is then calculated as follows:

$$\delta_p = \sum_g w_g \delta_g$$

The pooled standard error is calculated as below:

$$\sigma_p = \sqrt{\sum_g w_g^2 \sigma_g^2}$$

Appendix F. Descriptive Analysis of Student Outcome Data for ACE Program Participants and Non-participants

Table F1. Percentage of Participants and Non-participants Meeting Reading Standard by Grade, 2011

	Percentage Met TAKS-ELA/Reading Standard							
	30+ Days of Participation				60+ Days of Participation			
Grade	Treatment	N	Control	N	Treatment	N	Control	N
4	75.8%	9383	78.0%	20334	76.1%	6857	79.1%	25471
5	77.8%	8997	78.3%	24943	78.7%	6333	78.8%	23994
6	75.5%	10983	77.7%	61652	77.1%	6598	71.0%	8733
7	77.2%	9786	70.9%	11759	77.6%	5185	72.5%	10839
8	81.6%	8801	75.7%	11671	81.2%	4606	82.6%	45971
9	79.0%	4217	74.7%	28245	80.2%	1662	77.2%	39734
10	85.6%	3498	83.6%	21205	86.7%	1701	85.7%	41686
11	90.3%	3550	85.1%	4386	91.8%	1435	89.0%	5387

Source: TEA TAKS data, 2010–11

Table F2. Percentage of Participants and Non-Participants Meeting Math Standard by Grade, 2011.

	Percentage Met TAKS Mathematics Standard							
	30+ Days of Participation				60+ Days of Participation			
Grade	Treatment	N	Control	N	Treatment	N	Control	N
4	81.0%	9383	82.8%	20334	81.6%	6857	83.6%	25471
5	77.0%	8997	77.3%	24943	78.1%	6333	78.1%	23994
6	74.7%	10983	76.4%	61652	77.0%	6598	70.2%	8733
7	73.0%	9786	65.9%	11762	73.6%	5185	67.3%	10841
8	71.1%	8799	63.6%	11682	71.5%	4606	72.2%	45983
9	60.5%	4193	54.6%	28019	61.4%	1660	57.0%	39461
10	68.3%	3928	65.3%	21096	68.5%	1695	68.1%	41485
11	85.6%	3533	78.7%	3379	85.7%	1424	82.8%	5341

Source: TEA TAKS data, 2010–11

Table F3. Percentage of Participants and Non-participants Promoted by Grade, 2012

	Percentage Promoted in 2012							
	30+ Days of Participation				60+ Days of Participation			
Grade	Treatment	N	Control	N	Treatment	N	Control	N
4	96.6%	9510	89.7%	21538	97.0%	6927	92.4%	26644
5	96.1%	9073	91.0%	26082	96.7%	6368	91.8%	24996
6	97.3%	11065	94.0%	63264	97.5%	6630	90.0%	9267
7	96.8%	9901	85.4%	12746	97.4%	5224	90.1%	11557
8	96.2%	8905	87.2%	12656	96.7%	4649	92.9%	48213
9	87.8%	4289	74.1%	31074	90.5%	1683	77.6%	42814
10	91.4%	4026	82.3%	23088	94.2%	1722	85.7%	44196
11	87.7%	3924	70.4%	4386	90.1%	1553	76.6%	6603

Source: TEA PEIMS data, 2010–11

Table F4. Average Days Absent for Participants and Non-participants by Grade, 2011

	Average Days Absent, 2011							
	30+ Days of Participation				60+ Days of Participation			
Grade	Treatment	N	Control	N	Treatment	N	Control	N
4	4.82	9510	4.62	21538	4.51	6927	4.50	26644
5	4.82	9073	4.78	26082	4.40	6368	4.57	24996
6	5.93	11065	6.77	63264	5.21	6630	6.86	9267
7	6.14	9901	7.69	12746	5.60	5224	7.74	11557
8	6.60	8905	8.31	12656	6.03	4649	7.30	48213
9	8.33	4289	11.19	31074	7.31	1683	10.75	42814
10	8.58	4026	10.81	23088	7.40	1722	10.20	44196
11	9.29	3924	12.29	4386	8.50	1553	10.88	6603
12	10.53	3237	12.82	3390	8.79	1222	11.17	33730

Source: TEA PEIMS data, 2010–11

Table F5. Average Disciplinary Incidents for Participants and Non-participants by Grade, 2011

	Average Disciplinary Incidents, 2011							
	30+ Days of Participation				60+ Days of Participation			
Grade	Treatment	N	Control	N	Treatment	N	Control	N
4	0.20	9510	0.15	21538	0.18	6927	0.14	26644
5	0.32	9073	0.26	26082	0.28	6368	0.25	24996
6	1.03	11065	0.94	63264	0.82	6630	1.02	9267
7	1.20	9901	1.24	12746	1.00	5224	1.25	11557
8	1.14	8905	1.27	12656	0.96	4649	1.10	48213
9	0.97	4289	1.36	31074	0.88	1683	1.42	42814
10	0.74	4026	0.97	23088	0.64	1722	1.07	44196
11	0.54	3924	0.67	4386	0.40	1553	0.57	6603
12	0.34	3237	0.45	3390	0.30	1222	0.48	33730

Source: TEA PEIMS data, 2010–11