



Exploring the Relationship Between Afterschool Program Quality and Youth Development Outcomes: Findings From the Washington Quality to Youth Outcomes Study

Final Report

Neil Naftzger
Samantha Sniegowski

FEBRUARY 2018

Exploring the Relationship Between Afterschool Program Quality and Youth Development Outcomes: Findings From the Washington Quality to Youth Outcomes Study Final Report

February 2018

Neil Naftzger
Samantha Sniegowski



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

www.air.org

Copyright © 2018 American Institutes for Research. All rights reserved.

3690_02/18

Contents

	Page
Executive Summary	i
Major Study Takeaways	vi
Introduction.....	1
Conceptual Framework.....	2
Research Questions and Study Design	4
Afterschool Programs Enrolled in the Study	5
Summary of Study Data and Key Program Differences.....	6
Process Quality	7
Youth Engagement	11
Daily Activities	13
Youth Development Outcomes.....	16
Youth Summative Experiences in Programming.....	20
Youth Characteristics.....	23
Key Study Findings.....	24
Process Quality and Summative Youth Experiences Related to End-of-Day Engagement	24
Daily Activities Related to Youth Engagement.....	27
Changes in Youth Development Outcomes.....	28
Process Quality, Engagement, and Changes in Youth Development Outcomes.....	29
Group Discussion and Changes in Youth Development Outcomes	36
Summary of Key Findings	37
An Alternate Interpretation.....	38
Limitations	41
Conclusions.....	43
Next Steps.....	45
References.....	46

Tables

	Page
Table 1. Percentage of Pre-Post Survey Respondents With Room to Improve by Survey Scale.....	19
Table 2. Percentage of Postsurvey Respondents Indicating a Particular Program Impact	20
Table 3. Demographic Characteristics Associated With Youth Enrolled in the Study	23
Table 4. Demographic and Process Quality Predictors of Average Youth Engagement.....	25
Table 5. YPQA Subscales Used as Predictors of Average Youth Engagement	26
Table 6. Demographic and Process Quality Predictors of Average Youth Engagement With Grade-Level Status of the Center.....	27
Table 7. Daily Activities Predicting Youth Engagement in Programming	28
Table 8. Changes in Average Pre-Post Youth Survey Scores by Construct and Program Attendance Level	29
Table 9. Demographic, Process Quality, and Youth Experience Predictors of Changes in the Academic Identity Scale	31
Table 10. Demographic, Process Quality, and Youth Experience Predictors of Changes in the Positive Mindsets Scale	32
Table 11. Summary of Significant Cross-Level Interactions Between YPQA Scores and Youth-Level Predictors—Changes in the Positive Mindsets Scale	33
Table 12. Demographic, Process Quality, and Youth Experience Predictors of Changes in the Self-Management Scale	34
Table 13. Summary of Significant Cross-Level Interactions Between YPQA Scores and Youth-Level Predictors—Changes in the Self-Management Scale.....	34
Table 14. Demographic, Process Quality, and Youth Experience Predictors of Changes in the Interpersonal Skills Scale.....	34
Table 15. Summary of Significant Cross-Level Interactions Between YPQA Scores and Youth-Level Predictors—Changes in the Interpersonal Skills Scale	35
Table 16. Demographic, Activity, and Youth Experience Predictors of Changes in the Interpersonal Skills Scale.....	37
Table 17. Demographic and Mean Youth Engagement Predictors of the Youth-Reported Opportunities for Agency Scale.....	40
Table 18. Demographic and Mean Youth Engagement Predictors of the Perceptions of Other Youth Scale.....	40

Figures

	Page
Figure 1. Conceptual Framework	2
Figure 2. <i>Collaboration</i> and opportunities for <i>agency</i> were more prevalent in <i>secondary</i> programs.	10
Figure 3. There was a <i>wider range</i> of overall process quality among <i>elementary</i> programs than secondary programs as defined by the YPQA.	10
Figure 4. <i>Most</i> youth reported <i>high levels</i> of <i>engagement</i> during program activities, although this was especially the case for youth in secondary programs.....	12
Figure 5. The degree to which <i>engagement</i> scores across days varied by program.	13
Figure 6. For both elementary and secondary programs, working <i>alone</i> , working in <i>small groups</i> , and exploring <i>new things</i> were the <i>most common</i> things youth did during activities.	15
Figure 7. More than <i>80%</i> of the 282 youth with presurvey and postsurvey data had a baseline score that put them in either the <i>Mostly true</i> or <i>Completely true</i> range for each scale.....	18
Figure 8. More than <i>80%</i> of youth fell within the <i>Mostly true</i> or <i>Completely true</i> range of the scale for perceptions of activity leaders, while there was more variation in how youth viewed their peers in the program.....	21
Figure 9. The vast <i>majority</i> of youth provided responses that put them in the <i>Rarely/Sometimes</i> portion of the scale.....	22
Figure 10. Conceptual Framework	39
Figure 11. Adjusted Conceptual Framework – Focus on Measurement Approaches.....	41

Executive Summary

During the course of the 2016–17 school year, researchers from American Institutes for Research (AIR) undertook a study oriented at exploring the relationship between afterschool program *quality*, youth *engagement* in programming, and changes in a series of *youth development-related outcomes* measured through the *Youth Motivation, Engagement, and Beliefs Survey*. Funded by the Raikes Foundation and the Washington Office of Superintendent of Public Instruction (OSPI), the study was undertaken in 11 afterschool programs funded by the 21st Century Community Learning Centers (21st CCLC) program and targeted youth in Grades 4–9. The study was designed to answer three primary research questions:

1. Is there a relationship between youth-reported daily experiences in afterschool programming and (a) observed levels of program quality and (b) the types of learning opportunities undertaken on a given day?

A major goal of the study was to sample the experiences youth were having in the afterschool programs enrolled in the study and explore how these experiences were related to the level of program quality observed using the Youth Program Quality Assessment (YPQA). More specifically, we were interested in exploring if observed quality was high, did youth report being more engaged in program activities.

What data did we collect to answer this research question? To collect youth engagement data, we administered a paper youth experience survey at the end of afterschool programming that youth completed on two days a week during one week a month for up to six months during the 2016–17 school year. The research team designed this approach to obtain relatively immediate reactions from youth about the afterschool programming they had just participated in on the day in question. We collected 1,625 youth experience surveys between October 2016 and June 2017 from 440 individual youth enrolled in study programs.

Four items from these surveys were used to create an engagement measure: (1) *Were today's activities interesting?* (2) *Were today's activities important to you?* (3) *Did you enjoy today's activities?* and (4) *Did you have to concentrate to do today's activities?* Higher scores on the measure created from these items were interpreted as youth being more *engaged* in the afterschool activities they had just participated in on the day in question.

In addition to the end-of-day youth experience surveys, youth enrolled in study programs answered a series of questions on a postsurvey administered toward the end of the school year to obtain a more summative portrait of what youth experienced while participating in programs. More specifically, youth answered questions about how they felt about other youth attending the program and the activity leaders providing afterschool programming, and how frequently they participated in activities that may have afforded them the opportunity to experience a sense of agency.

We also were interested in exploring whether certain types of learning activities and opportunities were found to be associated with higher levels of youth-reported engagement in programming. On the days youth experience surveys were collected, the activity leaders responsible for the design and delivery of activities taking place on those days also

completed a survey of their own. The purpose of this survey was to obtain information on what youth did during a given activity, providing the research team with the capacity to explore how differences in what activities were provided on a given day may be related to youth-reported levels of engagement in programming. A total of 342 surveys were collected from activity leaders.

Finally, observations were conducted in spring 2017 to observe programming using a modified version of the YPQA to assess the level of program quality at each of the study programs. In addition to an overall mean total score, data we collected from the spring observations were used to calculate quality scores across four subscales unique to this study: (1) *Positive Emotional Climate*, (2) *Learning Supports and Format*, (3) *Collaboration*, and (4) *Opportunities for Youth Agency*. The YPQA was scored for a total 43 segments across the 11 programs enrolled in the study.

What did we find? A series of multilevel models were run to explore the relationship between program quality and the average level of youth-reported engagement in programming. In addition to data drawn from the sources described in the previous section, youth-level demographic data obtained from OSPI data warehouses were included in these analyses. As hypothesized, YPQA scores were found to be significantly related to the mean level of youth engagement, with higher scores associated with higher levels of engagement. This particularly was found to be the case in relation to the *Opportunities for Youth Agency* subscale of the YPQA, which was found to be especially related to mean engagement.

In addition, the degree to which youth directly reported on the postsurvey as having frequent opportunities to participate in activities related to supporting youth agency was found to be significantly and positively related to engagement. **In this sense, there was a consistent finding both from activity observations and youth reports that the more youth had opportunities to experience a sense of agency while participating in afterschool activities, the more apt they were to report being engaged in programming.**

In addition, certain activities also were found to be related to engagement. When youth participated in whole-group discussions led by the activity leader or listened to a presentation from a speaker or special guest from outside the program, they also reported being more engaged in programming.

2. To what extent do youth demonstrate improvement on youth development outcomes after participating in afterschool programming?

A growing body of research has shown that more learning time spent in high-quality afterschool programs helps young people develop the social and emotional skills and dispositions needed to be successful in school and in life (Durlak & Weissberg, 2007; Kauh, 2011; Miller, 2007; Pierce, Auger, & Vandell, 2013; Traill, Brohawn, & Caruso, 2013). However, what these studies have largely not done is specifically examine where youth start in terms of youth development-related skills, beliefs, and attitudes and what this means for how youth improve in these areas over time. We were particularly interested in assessing how youth that start with less developed skills and beliefs change as they participate in higher quality programming.

What data did we collect to answer this research question? Youth development-related outcomes were measured using the *Youth Motivation, Engagement, and Beliefs Survey*, a survey originally developed by the Youth Development Executives of King County (YDEKC) and modified by AIR to support statewide evaluation efforts associated with the Washington 21st CCLC program. The survey was administered on a pre-post basis at the start and end of the 2016–17 school year to assess the degree to which youth demonstrated growth in four areas: (1) *Academic Identity*, (2) *Positive Mindsets*, (3) *Self-Management*, and (4) *Interpersonal Skills*. A total of 282 youth completed the presurvey and the postsurvey.

What did we find? Paired sample *t* tests were run first to explore whether youth with room to grow on the survey outcomes based on their prescore demonstrated significant growth on the outcomes in question during the span of the school year. It is important to note that although only youth deemed to have room to improve on the outcome in question were included in these analyses, for the most part, most youth included in these analyses were already demonstrating largely positive functioning on the survey outcomes at the preadministrations of the survey. However, despite this finding, across each of the four constructs measured on the *Youth Motivation, Engagement, and Beliefs Survey*, youth demonstrated significant growth between administrations of the presurvey and the postsurvey.

Next, we were interested in exploring the degree to which program attendance was related to the degree of growth on the youth survey outcomes. Our hypothesis was that higher levels of program attendance would be associated with a greater degree of growth on the outcomes examined. Youth with pre-post survey results were classified into three groups based on their level of program attendance: (1) less than 30 days, (2) 30 to 59 days, and (3) 60 days or more. **We reran the analysis for each of these three groups and found that only youth in the 60-day or more attendance group demonstrated significant improvement in relation to the *Academic Identity, Positive Mindsets, and Self-Regulation* scales.** This finding seems to support our hypothesis that greater attendance in programming would be associated with higher levels of growth on the youth development outcomes examined.

3. Is there a relationship between program quality, youth-reported experiences in programming, and changes in key youth development outcomes?

Our goal here was to be able to document how quality practices may have led to youth having a set of experiences while participating in programming that cumulatively were associated with growth on key youth development outcomes. In this sense, we wanted to better describe those experiences youth have in programming that may bridge the relationship between quality practice and youth development outcomes.

What data did we collect to answer this research question? To answer this research question, we combined data from each of the sources described previously, including YPQA data, youth engagement data from the daily youth experience survey, summative youth experience data from the postsurveys, the daily activity leader surveys, youth demographic data, and data from the *Youth Motivation, Engagement, and Beliefs Survey*.

What did we find? A series of multilevel models were run to explore the relationship between program quality, youth experiences in programming, and growth on each of the four youth development-related outcomes measured on the *Youth Motivation, Engagement, and Beliefs Survey*. We hypothesized that both youth engagement and the level of program quality would be related to growth on the youth development-related outcomes measured by the survey.

Contrary to our hypothesis, we did not find that the mean level of youth-reported engagement was significantly related to improvement in the youth development outcomes assessed via the *Youth Motivation, Engagement, and Beliefs Survey*. Although this result was not expected, this finding may be related to the relatively sparse nature of our youth-level engagement data (youth included in these analyses provided an average of six daily youth experience surveys).

However, when we asked youth about their overall program experiences on the postsurvey, we found youth perceptions of their peers to be positively related to growth on each of the four youth development outcomes examined, suggesting the social climate of the program may be critical to moving the needle on these outcomes. In addition, we found the youth-reported opportunities for agency scale to be positively related to growth on the *Self-Management* and *Interpersonal Skills* scales (moderately significant) and especially on the *Positive Mindsets* scale. **In this sense, more positive relationships with program peers and greater opportunities for youth to experience a sense of agency were associated with greater improvement on the youth development-related outcomes measured by the pre-post survey.** As a result, considering the summative experiences of youth appeared to be important when exploring the extent to which youth had improved on the youth development outcomes under consideration.

We also found some evidence linking YPQA scores to changes in the youth development-related outcomes, but not in the manner initially hypothesized. Although we did not find any main effects linking YPQA scores to changes in the pre-post youth survey outcomes examined, we did find that select *YPQA scores seemed to strengthen* the positive relationship between youth-reported experiences in programming from the postsurvey and growth on three of the four youth outcomes (*Positive Mindsets*, *Self-Management*, and *Interpersonal Skills*). This largely occurred in relation to youth perceptions of (a) other youth attending programming in terms of how supportive and friendly their peers in the program were and (b) the opportunities they had to experience a sense of agency while participating in programming. **In this sense, higher YPQA scores strengthened the relationship between positive youth experiences in programming and growth on the youth development outcomes examined.**

Finally, the more center activity leaders reported offering activities characterized by *group discussions*, the more youth demonstrated improvement on the *Interpersonal Skills* scale of the pre-post youth survey, a finding that makes intuitive sense given the nature of the activity and the outcome in question.

An Alternate Interpretation

In undertaking the analyses outlined in the report, our hypothesis was that higher quality programs would lead to higher levels of youth engagement in programming on a day-to-day basis and that greater engagement over time would be associated with greater improvement on the youth development outcomes examined. Both YPQA scores and scores on the *Youth-Reported Opportunities for Agency* scale derived from responses from the post-youth survey were treated as measures of program quality. We hypothesized that higher scores on each of these measures would be associated with higher levels of youth-reported engagement in programming, a hypothesis that was found to be consistent with our results.

However, it is important to point out that youth engagement data was collected *prior* to post-youth survey data. In this sense, the degree to which youth experienced engagement on a day-to-day basis while participating in programming may have also influenced how they responded to the *Youth-Reported Opportunities for Agency* scale. A similar case can be made that day-to-day experiences in programming also likely influenced how youth responded to the *Perceptions of Other Youth* scale also contained on the post-survey, which asked a series of questions about how supportive and friendly other youth in the program were. In this sense, it may be justifiable to treat these postsurvey scales as outcomes that could be influenced by youth's day-to-day experiences in programming. When we ran models to test this possibility, we found that higher engagement scores were associated with higher scores on each of these postsurvey scales when the latter were treated as outcomes.

With these results, it seems like there may some preliminary evidence that suggests a slightly revised chain of events when trying to understand the relationship between afterschool program participation and youth outcomes akin to the following:

- Higher YPQA scores were associated with higher reported average engagement in programming on a day-to-day basis;
- Higher levels of engagement in day-to-day programming was found to be related to more positive youth descriptions of their summative and cumulative experiences in programming, both in terms of positive peer interactions and more frequent opportunities for agency;
- More positive youth-reported peer interactions and greater opportunities for agency were found to be associated with greater growth on the youth-development outcomes examined.

We find this sequence of significant relationships connecting program quality to positive youth experiences in programming to improvement on youth development outcomes to be of particular interest. If anything, this sequence of events may provide an initial template to support future measurement efforts to connect afterschool program quality to improvements on youth development outcomes assessed as part of this study.

Major Study Takeaways

One of the major takeaways from this study is the importance of accounting for individual youth experiences in programming when attempting to assess youth growth on outcomes akin to those assessed in our study and that higher program quality at the setting level can serve to strengthen the connection between youth experiences in programming and positive youth outcomes. In particular, both ensuring a positive social climate within the program where youth feel supported and valued by their peers and providing youth with opportunities to experience a sense of agency while participating in afterschool activities appeared especially related to both supporting youth engagement in programming and growth on the youth outcomes measured by the *Youth Motivation, Engagement, and Beliefs Survey*. Higher program quality, as measured by the version of the YPQA we used in the study, seemed to strengthen the relationship between youth-reported experiences in programming and growth on the youth development-related outcomes under consideration.

Introduction

During the span of the past several years, both the Office of Superintendent of Public Instruction (OSPI) and the Raikes Foundation have made significant investments in supporting the development of an infrastructure within the state of Washington to enhance the quality of afterschool programs that primarily serve youth living in higher poverty communities. Key to this strategy has been the adoption of the Youth Program Quality Assessment (YPQA), a validated observation instrument developed by the David P. Weikart Center for Youth Program Quality, which articulates what types of supports and opportunities should be available to youth participating in afterschool programs in order to support positive youth development. With support from OSPI and the Raikes Foundation, afterschool programs in Washington have been asked to use the YPQA as a program self-assessment tool in order to determine program delivery strengths and areas in which improvement may be warranted; craft action plans oriented at targeting specific areas for improvement; and access YPQA-related supports and trainings targeted at improving practitioner knowledge of and capacity to design and deliver developmentally appropriate learning environments for participating youth.

An initial body of evidence supports the potential efficacy of the investment being made by OSPI and the Raikes Foundation in terms of the link between program quality and youth outcomes. For example, studies completed by American Institutes for Research (AIR) have shown that higher levels of program quality as defined by the YPQA are related to (a) higher levels of youth engagement in programming; (b) sustained participation in afterschool programming over time; and (c) positive school-related outcomes, including academic achievement and fewer school-day disciplinary referrals (Naftzger, Devaney, & Foley, 2014; Naftzger, Hallberg, & Tang, 2014; Naftzger et al., 2013).

However, less is known about how *sustained* participation in high-quality afterschool programs as defined by the YPQA serves to impact students in terms of key youth development outcomes, including those related to the formation of positive mindsets and beliefs and social and emotional skills and competencies. It is hypothesized that sustained participation in high-quality programs will support youth growth and development in these areas, which in turn will contribute to larger school-related gains over time.

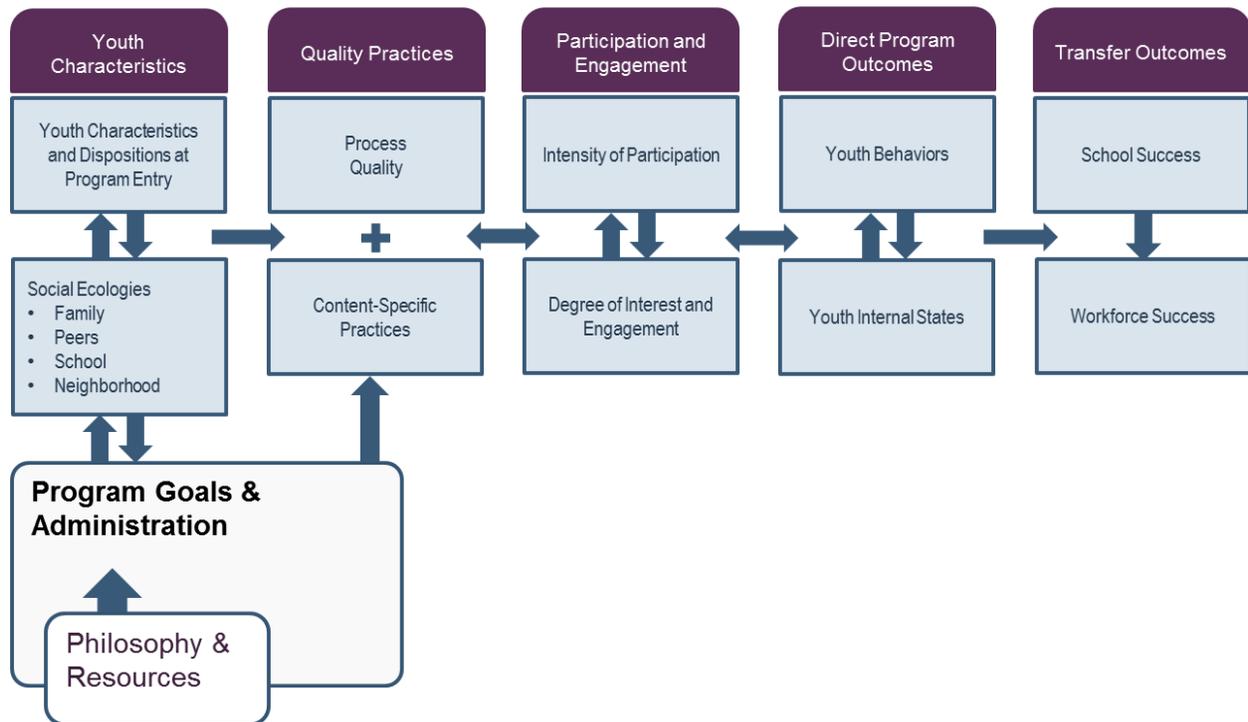
In order to explore this hypothesis, OSPI and the Raikes Foundation provided funding to AIR to conduct a one-year study to explore the relationship between youth experiences in a sample of afterschool programs funded by the 21st Century Community Learning Centers (21st CCLC) program and changes in youth perceptions of how they are functioning on key youth development outcomes, while controlling for the level of observed program quality. Dubbed the Washington Quality to Youth Outcomes Study, this effort was undertaken during the span of the 2016–17 school year in 11 21st CCLC-funded programs located across the state of Washington. The programs included in this study were also taking part in a statewide initiative called the *Expanded Learning Opportunity Quality Initiative*. This initiative was funded by a public/private partnership to pilot a quality system that seamlessly connects with the birth to five Quality Rating Improvement System (QRIS) in youth development programs. The 21st CCLC programs enrolled in the study represented a subset of the total pilot sites involved in the Initiative.

The purpose of this report is to outline key findings from the study and provide recommendations regarding what these results mean in terms of both further supporting the efforts of afterschool practitioners to design and deliver high-quality programs and efforts to measure the manner in which these programs may be contributing to positive youth outcomes.

Conceptual Framework

The quality-related infrastructure being cultivated by investments from OSPI and the Raikes Foundation is predicated on a conceptual framework similar to one developed by AIR regarding how youth benefit from participation in afterschool programs (see Figure 1). The framework starts with the youth themselves and how they are influenced and supported by the environments in which they live and go to school. Past programming experiences, relations with peers and teachers, the level of interest in programming topics and content, expectations regarding program experience, and the level of choice in attending all have a bearing on how youth will engage in and experience afterschool programming (Durlak, Mahoney, Bohnert, & Parente, 2010).

Figure 1. Conceptual Framework



A number of factors influence the experience youth have once they are in the program. First, programs must be of high quality to have an impact. Generally, there are two categories of quality: (1) process quality and (2) content-specific practices. Process quality refers to the adoption of practices and approaches to service delivery that result in the creation of a developmentally appropriate setting for youth, where participants feel safe and supported and are afforded opportunities to form meaningful relationships, experience belonging, and be an active participant in their own learning and development. These practices are universal because they are truly applicable to any type of youth programming, irrespective of content, approach, or setting. These are the type of practices described in the YPQA. In conducting the Washington Quality to Youth Outcomes Study,

we focus on this form of quality because it represents the primary area in which OSPI and the Raikes Foundation have sought to improve afterschool program quality in the state.

Content-specific practices are program practices designed to intentionally cultivate a specific set of skills, beliefs, or knowledge. Often, such practices are closely aligned with the direct outcomes a program is seeking to cultivate in participating youth. For example, content-specific practices include specific approaches to cultivating literacy skills, formal curricula for social and emotional learning, or methods of teaching technology skills. Content-specific practices adopted by afterschool programs are remarkably diverse and, therefore, are measured and assessed in a variety of different ways. For the Washington Quality to Youth Outcomes Study, we did not take steps to measure how programs were functioning on this component of quality.

Each of these two initial components (i.e., youth characteristics and quality practices) inform and are informed by both the philosophy of the program and the goals and objectives they set out to accomplish with participating youth. All of the programs enrolled in the study were funded by 21st CCLC and, therefore, were influenced by the school-related outcomes sought by this federal funding stream around promoting academic achievement and promoting positive school-related behaviors.

For youth to benefit from programming, they need to attend programming, ideally at high levels and in a variety of different types of activity. Being merely present in the program is not enough, however, to ensure that youth will benefit from activities. They need also to experience both engagement and interest during their activities in order to develop the beliefs, skills, and knowledge that can help them in school and beyond. In theory, the extent to which programs effectively adopt both practices related to process quality and content-specific practices should heavily influence the degree of engagement and interest youth experience while participating in 21st CCLC programming. Exploring the degree of youth engagement in programming was a substantial component of the Washington Quality to Youth Outcomes Study.

Once youth are engaged and participating, it is expected that they will begin to develop key skills, beliefs, and knowledge based on their participation in program activities. These are termed *direct program outcomes* in the conceptual framework outlined in Figure 1. Based on AIR's research into 21st CCLC programs over the past decade, direct program outcomes fall into two categories: (1) academic knowledge, attitudes, and behaviors and (2) social and emotional skills and beliefs (broadly defined), which can be demonstrated by both youth behaviors that are directly observable by others and internal states and beliefs that need to be elicited from youth in some way. Development of a viable and valid measurement strategy in relation to assessing how programs support development in relation to these outcomes is an almost existential need for the afterschool field in order to better demonstrate the manner in which quality programs are having an impact on youth. Being able to learn more about how measures focusing on documenting potential program impacts in these areas work was a major focus of the Washington Quality to Youth Outcomes Study.

Finally, the skills, beliefs, and knowledge youth develop through their participation in high-quality afterschool programming may be used in other settings outside of the program to drive achievement and success in school and the workplace. These outcomes typically are measured by connecting participation data with school-related data available at the state or local level. These outcomes were not assessed in the present study, largely because of time and resource constraints.

Overall, then, the Washington Quality to Youth Outcomes Study was focused on understanding the relationship between *process quality* instantiated by the YPQA, the degree of youth *engagement* reported by youth, and *youth self-assessments* of how they were functioning in a series of areas related to (1) attitudes and behaviors related to school and (2) social and emotional skills and beliefs.

Research Questions and Study Design

Building from the conceptual framework on how we see afterschool programs having a positive impact on youth, the Washington Quality to Youth Outcomes Study was designed to answer the following set of research questions.

1. *Is there a relationship between youth-reported daily experiences in afterschool programming and (a) observed levels of program quality and (b) the types of learning opportunities undertaken on a given day?* The goal here was to sample the experiences youth were having in afterschool programs over a defined period of time and validate the connection between the level of quality practice observed using the YPQA and positive youth experiences hypothesized to be related to the key youth development outcomes of interest. More specifically, we were interested in exploring whether observed quality was high and whether youth reported being more engaged in program activities. In addition, we also were interested in exploring whether certain types of learning activities and opportunities were found to be associated with higher levels of youth-reported engagement in programming.
2. *To what extent do youth demonstrate improvement on youth development outcomes after participating in afterschool programming?* A growing body of research has shown that more learning time spent in high-quality afterschool programs helps young people develop the social and emotional skills and dispositions needed to be successful in school and in life (Durlak & Weissberg, 2007; Kauh, 2011; Miller, 2007; Pierce, Auger, & Vandell, 2013; Trill, Brohawn, & Caruso, 2013). However, what these studies have largely not done is specifically examine where youth start in terms of youth development-related skills, beliefs, and attitudes and what this means for how youth improve in these areas over time. We were particularly interested in assessing how youth that start with less developed skills and beliefs change as they participate in higher quality programming.
3. *Is there a relationship between program quality, youth-reported experiences in programming, and changes in key youth development outcomes?* Our goal here was to be able to document how quality practices may have led to youth having a set of experiences while participating in programming that cumulatively were associated with growth on key youth development outcomes. In this sense, we wanted to better describe those experiences youth have in programming that may bridge the relationship between quality practice and youth development outcomes.

In answering this set of research questions, a descriptive, correlational study design was undertaken that focused on exploring the relationship between key constructs articulated in the research questions without seeking to establish a causal relationship between program quality, youth experiences in programming, and changes in youth development-related outcomes. Our intent was to create a body of results that can be used to further inform development and

refinement of quality measurement and program improvement strategies and efforts to create viable measurement strategies for assessing the domain of youth development-related outcomes hypothesized to be associated with afterschool program participation.

Afterschool Programs Enrolled in the Study

A total of 11 programs funded by the Washington 21st CCLC program were enrolled in the Washington Quality to Youth Outcomes Study. These programs were selected for inclusion in the study because they had also been enrolled in an intervention study also supported by OSPI and the Raikes Foundation called the *Expanded Learning Opportunity Quality Initiative* undertaken by Cultivate Learning at the University of Washington and School's Out Washington, an afterschool intermediary that has played an active and substantive role in helping to construct the Raikes-funded quality improvement infrastructure in Washington. The OST Quality Initiative was oriented at testing a coaching intervention designed to support program adoption of the YPQA and to help sites more effectively move through the quality improvement process once a self-assessment had been completed using the tool. The thinking was that the two studies could exchange data collected as part of each undertaking to enhance the robustness of each individual study.

Of the 11 programs enrolled in the Quality to Youth Outcomes Study,

- six were located in elementary schools, four in middle schools, and one in a high school;
- six were located in the Eastern portion of the state in Spokane and Walla Walla Counties, while five were located in the Western side of the state in King and Pierce Counties;
- eight were administered by the local school district, two by an educational service district, and one by a community-based organization
- all programs enrolled in the study were physically located in public schools; and
- the number of enrolled youth per program for the 2016–17 school year ranged from 60 to 452, with a median of 85 youth served per program.

Relative to the other programs funded by 21st CCLC in the state, this group was distinguished by a willingness to participate in both the *Expanded Learning Opportunity Quality Initiative* and the Washington Quality to Youth Outcomes Study, both time-consuming endeavors in terms of data collection and the dedication of staff time to participate in quality improvement processes. It is likely that this willingness made these programs potentially different in other ways from the full domain of programs funded by 21st CCLC statewide. For participating in these efforts, each program received a stipend of \$2,000 for being part of the *Quality Initiative* and \$1,500 for being part of the Quality to Youth Outcomes Study.

Within this report, some figures will provide program-level information about the sites enrolled in the study. Pseudonyms have been used to maintain site confidentiality.

Summary of Study Data and Key Program Differences

To address the research questions underpinning the study, a variety of data collection activities happened during the span of the 2016–17 school year at the 11 21st CCLC-funded programs enrolled in the study. Below, we offer a brief introduction to each of these activities along with a brief preview of what we learned from these data relative to the programs enrolled in the study and the youth they served.

- *Process Quality*—Process quality was assessed by conducting observations of each of the study sites using a modified version of the YPQA in spring 2017. We found some notable differences in YPQA scores between elementary and secondary programs. Specifically, secondary programs scored higher on scales that described staff instructional practices that afforded youth a greater opportunity to *collaborate* with their peers and experience a *sense of agency* while participating in programming. In addition, the range of quality across programs was found to be broader across the elementary sites enrolled in the study as compared to the secondary sites, where the difference in quality across programs was smaller.
- *Youth Engagement*—Youth engagement was measured through administration of a print copy survey at the end of the programming day for up to 12 randomly selected programming days during the course of the 2016–17 school year. Here again, we found differences in scores between elementary and secondary programs. Youth at secondary programs had a tendency to demonstrate higher youth engagement scores than youth attending elementary programs. However, more variation in scores was found across individual days when surveys were administered for a given program as compared to the mean engagement score across programs. In this sense, the level of youth-reported engagement had a tendency to fluctuate from one day to the next, although this variation was more extensive in some programs than in others.
- *Daily Activities*—On the days that youth engagement data were collected, activity leaders also were asked to complete a survey for *each activity* they provided on that day. This survey asked questions related to *how* youth spent their time during the activity and how much time they spent doing it. For example, categories like *Youth worked on a group project that will take multiple sessions to complete* and *Youth learned or practiced a skill that is not related to a specific school-day content area (e.g., learning tae kwon do, etc.)* were included on the survey. Some differences in the degree to which certain activities happened were also noted between elementary and secondary programs. Specifically, activities more prevalent in secondary programs afforded youth a greater degree of autonomy and opportunities for agency while participating in programming.
- *Youth Development Outcomes*—Youth development outcomes were measured through the pre-post administration of the *Youth Motivation, Engagement, and Beliefs Survey* during the span of the 2016–17 school year. Changes in four youth development outcomes were measured on the survey: (1) *Academic Identity*, (2) *Positive Mindsets*, (3) *Self-Management*, and (4) *Interpersonal Skills*. We were particularly interested in assessing growth on these outcomes only for youth who had room to improve based on their presurvey score. This ranged from 43% to 60% of youth taking the presurvey and the postsurvey, depending upon on the outcome in question. Unfortunately, those that did have room to grow were largely already functioning at a positive level on these outcomes

at baseline, raising concerns about our ability to meaningfully detect growth on these scales. Youth in secondary programs were more apt to have room to grow on these outcomes compared to youth in elementary centers.

Youth were also directly asked to indicate how they thought they had benefited from program participation. Many of the top areas identified by youth had to do with the development of interests in new areas as well as the development of a positive self-concept and new friendships.

- *Youth Summative Experiences in Programming*—On the postadministration of the *Youth Motivation, Engagement, and Beliefs Survey*, youth also were asked to report on their perceptions of activity leaders and their peers in the program and the extent to which they had the opportunity to participate in activities likely to cultivate a sense of agency. Although we found a relatively small degree of variation between programs on these scales, there were notable differences across individual youth. In this sense, youth subjective experiences in programming appear to be quite different from one youth to the next and, therefore, are important to consider when assessing how youth may have benefited from participation in programming.
- *Youth Characteristics*—Data on youth demographics characteristics were obtained from OSPI directly. Youth enrolled in the study were in Grades 4–9 during the course of the 2016–17 school year and largely were representative of the larger 21st CCLC population in Washington (e.g., being eligible for free and reduced-price lunches, ethnic minorities, and scoring below proficiency on the statewide assessment in reading and/or mathematics).

A more detailed explanation of each data source can be found in the following sections of the report.

Process Quality

The YPQA was used to assess process quality at the 11 21st CCLC programs enrolled in the study. However, the measure used by the research team from AIR deviated from the normal YPQA tool, both in terms of the items appearing on the measure and the manner in which activities were scored. The version of the YPQA observation tool commonly used to support program self-assessment and external observations in Washington contains a total of 62 items, organized around four broad domains: (1) safe environment, (2) supportive environment, (3) interaction, and (4) engagement. Items on the YPQA are typically averaged to create mean scores for each of these four domains and one overall total score. In addition, when YPQA observation are completed, individual afterschool offerings typically are observed from beginning to end on the day the observation is conducted.

Our goal was to create a shorter version of the tool that could be used to observe shorter activity segments (approximately 30 minutes in length). This allowed observers to see more activities offered by a given site on the day of the visit, resulting in a mean YPQA score based on a greater breadth of programming offered by the site. This approach was adopted because we wanted a *program-level* measure of quality that took into consideration the greatest number of activities possible to better represent the variety of experiences youth may have while participating in

programming. Twenty-four items were selected for inclusion in the final tool, drawn from the YPQA, the STEM PQA, and the Academic Skill Building PQA.¹ Items chosen for inclusion on the tool were based on two considerations: (1) the item was shown in past studies completed by AIR to demonstrate significant variation across observations and (2) the items were hypothesized to provide the types of opportunities and supports that may lead to both youth engagement in programming and growth on the outcomes we attempted to measure as part of the study.

We used Rasch analysis² techniques to combine the domain of items into four subscales characterized by good reliability and functioning on other psychometric criteria. The 1, 3, and 5 rating scale associated with the full YPQA was retained, with a 5 indicating that a given support or opportunity described by the item was afforded to youth at the highest quality level, while a 1 indicated that a practice was largely absent, indicating low quality. The revised scales and items represented in each scale are outlined below:

1. Positive Emotional Climate (3 items)
 - Staff encourages all youth to try out skills or attempt higher levels of performance
 - The emotional climate of the session is predominantly positive ...
 - Youth do not exhibit any exclusion or staff successfully intervenes if exclusive behavior occurs ...

2. Learning Supports and Format (10 items)
 - Staff explains all activities clearly ...
 - The activities balance concrete experiences, materials, people, and projects with abstract learning and concepts ...
 - More than once, staff help youth connect current activity to personal experiences, applications, or previous knowledge ...
 - Activities are appropriately challenging for all or nearly all the participants ...
 - Staff models skills for all youth
 - When youth struggle, staff always provides learning supports or encouragement ...
 - There are two or more instances with different individual youth in which staff-youth conversations include substantive back-and-forth dialogue about offering content ...
 - Staff supports at least some contributions or accomplishments of youth by acknowledging what they have said or done with specific, nonevaluative language ...
 - Staff makes frequent use of open-ended questions ...
 - Staff is almost always actively involved with youth ...

¹ The STEM PQA and Academic Skill Building PQA are variations of the YPQA that contain additional items pertinent to assessing programs that involve the delivery of STEM content and content oriented at supporting youth academic skill building.

² Rasch analysis is an approach to modeling based on a latent trait that takes into consideration the interplay between the difficulty of the item appearing on the assessment and, in this case, the observed quality of the setting.

3. Collaboration (3 items)

- Staff provides opportunities for all youth to work cooperatively as a team or in a group
- Staff provides all youth opportunities to participate in activities with interdependent roles ...
- Staff provides opportunities for all youth to work toward shared goals

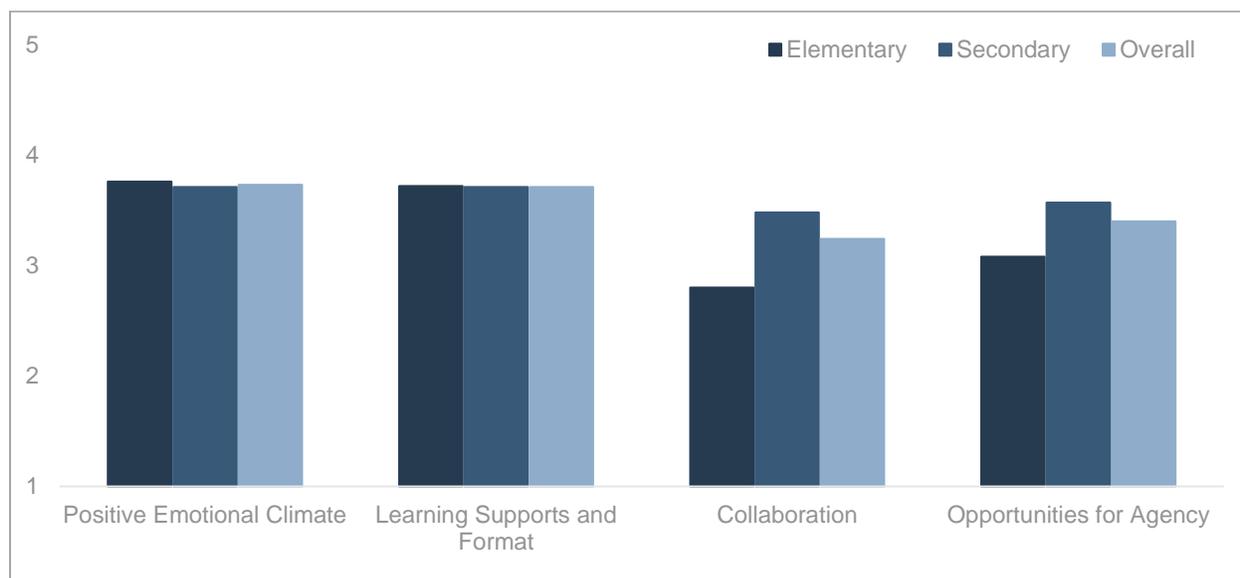
4. Opportunities for Youth Agency (8 items)

- The activities involve youth engaging with materials or ideas or improving a skill through guided practice for at least half of the time
- During activities, staff provides all youth a structured opportunity to talk about what they are doing and what they are thinking about to others ...
- The program activities lead to tangible products or performances that reflect ideas or designs of youth ...
- Staff attributes success to effort, attention, practice, or persistence ...
- Staff shares control of most activities with youth, providing guidance and facilitation while retaining overall responsibility ...
- Staff provides multiple opportunities for youth to make plans for projects and activities ...
- Staff provides opportunities for all youth to make at least one open-ended content choice within the content framework of the activities ...
- Staff provides opportunities for all youth to make at least one open-ended process choice ...

A total of 43 segments lasting approximately 30 minutes each were observed at the 11 sites between March and May 2017. Program observations took place during either a one- or two-day site visit conducted by consultants hired and trained by School's Out Washington. The number of segments observed per site ranged from two to seven, with an average score of 3.91. Generally, a larger number of observation segments occurred at the middle and high school sites, given a greater number of different activities and staff leading those activities. In this sense, more observation segments were needed in order to obtain a relatively comprehensive assessment of what was taking place at the center from a programming perspective.

Mean scores by subscale are shown in Figure 2, broken down by programs serving youth in elementary and secondary grade levels. Means scores exceeded 3.70 both overall and for programs serving both grade-level types on the *Positive Emotional Climate* and *Learning Supports and Format* scales; however, secondary centers demonstrated substantially higher scores on the *Collaboration* and *Opportunities for Agency* scales than elementary centers.

Figure 2. Collaboration and opportunities for agency were more prevalent in secondary programs.



Note: Based on 43 segment observations conducted at 11 study centers in spring 2017

This is not surprising given that the enhanced cognitive abilities that come with early adolescence allow youth to play a more active role in pursuing goals and objectives akin to those commonly associated with afterschool programming, such as working on a group project that takes several sessions to complete.

We also averaged the four subscale scores to create mean total YPQA score for each program. As shown in Figure 3, there was more variation in mean total YPQA scores across the elementary programs enrolled in the study as compared to the secondary programs.

Figure 3. There was a wider range of overall process quality among elementary programs than secondary programs as defined by the YPQA.



Note: Based on 43 segment observations conducted at 11 study centers in spring 2017

As described later in the report, we used the YPQA scores for each program in models exploring the relationship between program quality and youth-reported levels of engagement in programming *and* changes in perceived functioning on a series youth development-related outcomes.

Youth Engagement

One key characteristic of the Washington Quality to Youth Outcomes Study was the collection of extensive data from each program on the experiences youth had while participating in afterschool activities. We did this through the administration of a paper youth experience survey that youth completed on two days a week during one week a month for up to six months. The research team randomly selected the days the surveys were administered, prepared packets of print copy surveys, and sent the packets to study programs one week before data collection was to take place. Program staff were responsible for administering the survey at the end of programming for a selected day and then returned the completed surveys to the research team at AIR. Surveys were coded in order to track individual youth responses over time.

The research team designed this approach to obtain relatively immediate reactions from youth about the afterschool programming they had just participated in on the day in question. There were two key advantages to this approach: (1) Youth reported on recent events and experiences, thereby enhancing the quality and authenticity of their responses given less difficulty with recall; and (2) we were able to capture experiences from youth who both persisted in participating in afterschool programming at the site and those whose participation was less consistent in programming.

Our goal was collect youth experience surveys from study programs on 12 different programming days taking place during six weeks in the 2016–17 school year. This threshold was met with seven of the 11 study programs, with the number of days in which youth experience data was collected ranging from seven to 13. Reasons days were missed included program closure because of inclement weather, field trips where surveying was less viable, site coordinator absence on the days surveys were to be collected, and site coordinator forgetfulness or confusion around when exactly surveys were to be administered. We collected 1,625 youth experience surveys between October 2016 and June 2017, which accounted for an average of approximately 148 surveys per site, and ranged from 68 to 249 surveys per program. These 1,625 surveys represented 440 youth overall and ranged from 14 to 75 youth per program.

Programs serving elementary youth had more youth experience surveys completed on average (ranging from 4.12 to 8.11 surveys per youth) as compared to the secondary programs (which ranged from 1.80 to 3.83 surveys per youth). This difference reflects a contrast in service delivery models between the elementary and secondary programs. In the case of the latter, programming was typically delivered in a club format, where different activities were offered on different days and youth were apt to attend the clubs they were most interested in. At the elementary level, youth were more apt to be enrolled in the program and attend on a daily basis, participating in whatever activities may have been offered on the day in question.

For the purposes of the Washington Quality to Youth Outcomes Study, we were specifically interested in measuring youth *engagement* in the afterschool programming provided on a given day. Based on a definition of engagement used by Shernoff and Vandell (2007) in a similar study of youth participation in afterschool programs, engagement for our study was defined as

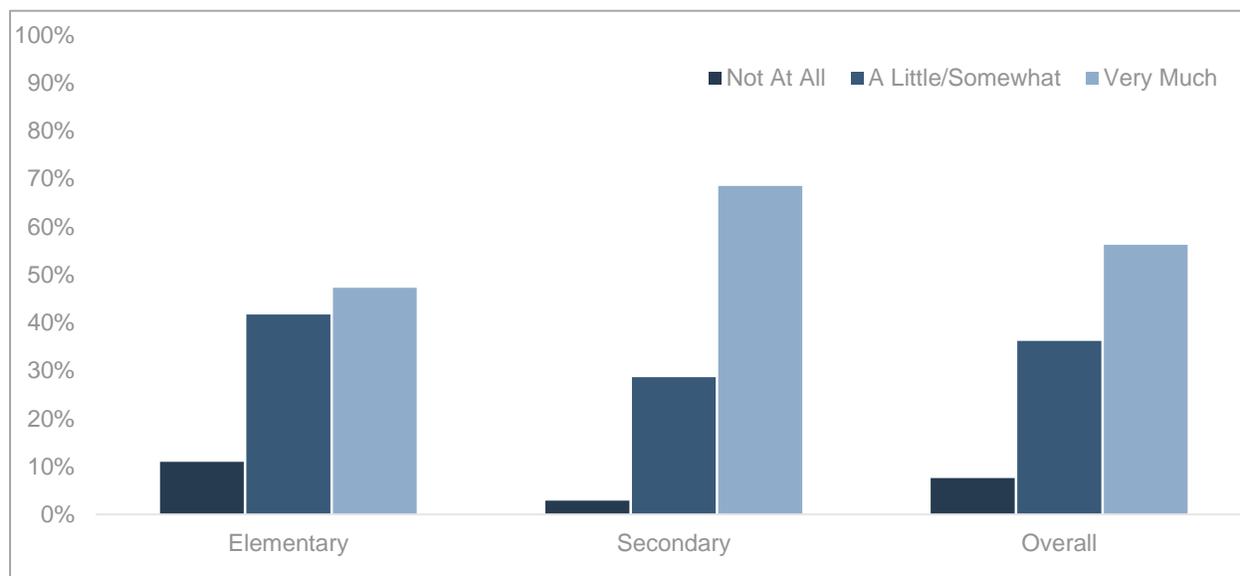
occurring when youth reported simultaneously experiencing *interest* in afterschool activities, *concentrating* while participating, *enjoying* the activities, and feeling that the activities were *important*. The last construct, feeling that activities were important, was not part of the engagement construct initially conceptualized by Shernoff and Vandell but was added to the scale in our study based on work being done in a parallel study by AIR, which demonstrated the connection between the concept of perceived activity importance and youth-reported interest development in STEM-oriented summer learning programs. Overall then, the engagement scale was based on the following four items ($\alpha = .85$):

1. Were today’s activities interesting?
2. Were today’s activities important to you?
3. Did you enjoy today’s activities?
4. Did you have to concentrate to do today’s activities?

In answering these questions, youth responded to a 4-point rating scale: (1) Not at all, (2) A little, (3) Somewhat, and (4) Very much. Rasch analysis techniques were used to create the engagement scale. As part of the scale creation process, it was determined that responding youth had some difficulty distinguishing between the *A little* and *Somewhat* response categories, so these response categories were collapsed when calibrating the final engagement measure.

As shown in Figure 4, the majority of youth reported being *Very much* engaged in program activities, although a substantive difference was found to exist between elementary and secondary programs. While 69% of youth in secondary programs responded *Very much* when asked questions about their level of engagement in programming, only 47% of youth in elementary programs provided responses at this level.

Figure 4. Most youth reported high levels of engagement during program activities, although this was especially the case for youth in secondary programs.

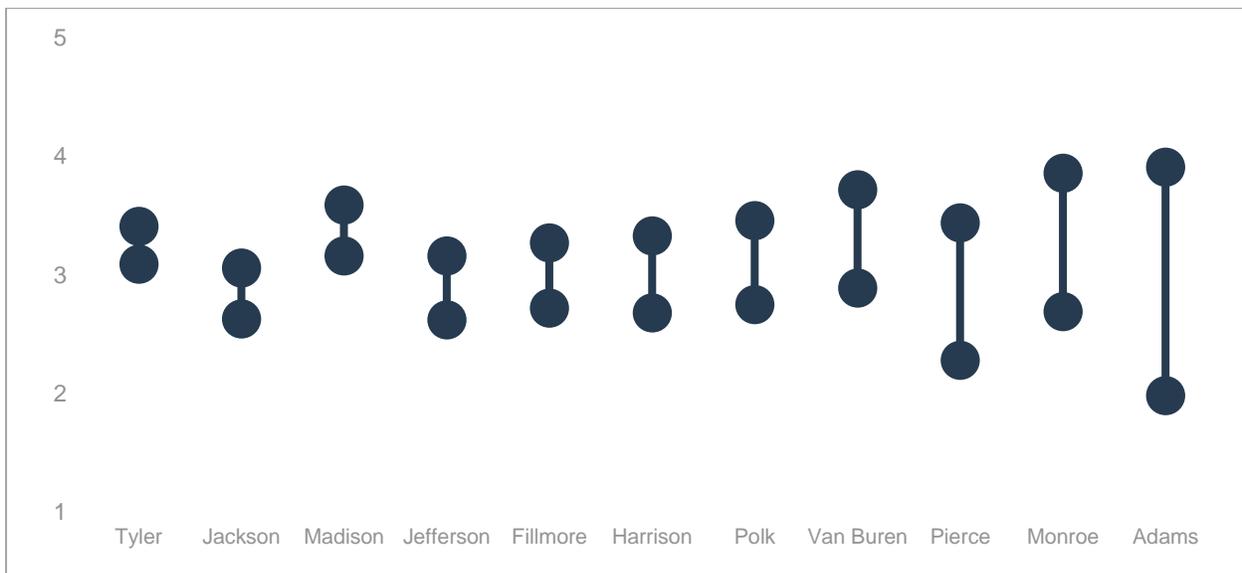


Note: Based on 1,625 youth experience surveys collected from the 11 study centers during the 2016–17 school year

We hypothesize that this difference may be related to the degree of choice youth had in selecting and attending activities that they were interested in, with youth in secondary programs having more options in this regard given the club format adopted by each of the secondary programs. However, the spread of overall mean engagement scores across programs within a given grade-level type was not found to be terribly broad, ranging from 2.85 to 3.16 for elementary centers and 3.22 to 3.34 for secondary centers.

However, as shown in Figure 5, more variation in scores was found across individual days when surveys were administered for a given center. This was particularly the case at three centers—Adams (secondary), Monroe (secondary), and Pierce (elementary). For example, for Adams, engagement scores ranged from 1.98 to 3.91 depending on the day engagement data was collected. In this sense, the “goodness of fit” between what programs offered in the way of activities and the degree to which youth were engaged in program offerings was found to vary from one day to the next. No center enrolled in the study offered the same activities day after day in a given week, so this variation is likely representative of the variation in activities from one day to the next, and in the case of secondary centers, different youth attending those activities. In this sense, the activities taking place on a given day then seems particularly important for some sites when exploring youth engagement.

Figure 5. The degree to which *engagement* scores across days varied by program.



Note: Based on 1,625 youth experience surveys collected from the 11 study centers during the 2016–17 school year

Daily Activities

In addition to collecting youth experience survey data on up to 12 afterschool programming days during the 2016–17 school year, the activity leaders responsible for the design and delivery of activities taking place on the days youth experience data were collected also completed a survey of their own. The purpose of this survey was to obtain information on what youth did during a given activity, providing the research team with the capacity to explore how differences in what activities were provided on a given day may be related to youth-reported levels in engagement in

programming. We asked activity leaders to report on whether youth did the following types of things during the activity:

- Youth primarily worked alone on tasks related to the activity (Alone)
- Youth primarily worked in small groups on tasks related to the activity (Small Groups)
- Youth received *direct instruction*³ in a particular academic content area (e.g., mathematics, science, reading, etc.) (Direct Instruction)
- Youth worked on a project that required them to make or build things (Build Things)
- Youth worked on a group project that will take multiple sessions to complete (Multiple Sessions)
- Youth participated in activities that allowed them to explore and discover new things on their own (New Things)
- Youth learned or practiced a skill that is not related to a specific school-day content area (e.g., learning tae kwon do, etc.) (Non-Academic Skill)
- Youth participated in a competition, contest, or game (Competition)
- Youth participated in whole-group discussions you facilitated (Group Discussion)
- Youth delivered a presentation to the whole group or an external audience (Give Presentation)
- Youth went on a field trip (Field Trip)
- Youth listened to a presentation from a speaker or special guest from outside the program (Guest Speaker)
- Youth planned future activities or projects (Planning)
- Youth participated in an activity that was designed to make a contribution or be helpful to others or the community (e.g., service learning project) (Service Learning)

For each of the aforementioned activities, activity leaders selected one of the following three options to indicate how prevalent the activity was during the session delivered on the day in question: (1) *No programming time was spent doing this*, (2) *Less than half of the programming time today was spent doing this*, and (3) *A majority of programming time today was spent doing this*.

Similar to the youth experience surveys, activity leader surveys were collected via print copy surveys sent to study sites a week before data collection was slated to be undertaken. Activity leaders completed a separate survey for each activity they delivered on the day in question and in which youth enrolled in study participated (i.e., activity leaders might have delivered multiple activities on a given day and only one of them might have targeted youth enrolled in the study).

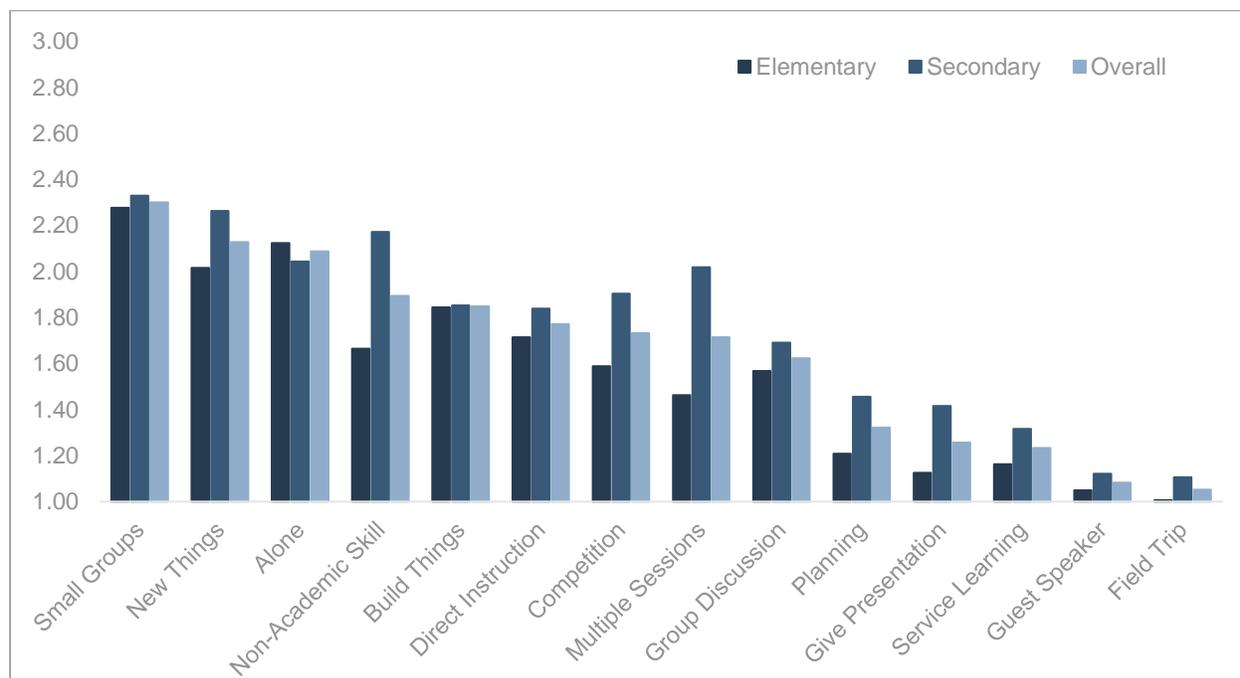
A complete set of activity leader surveys were returned for a median of nine days per site, ranging from four to 12 days, resulting in a total of 342 useable activity leader surveys. The number of

³ The term *direct instruction* referred to the following set of activities: (1) *Group Instruction*—These activities largely mirror typical school-day classroom instruction with the activity leader or teacher spending the bulk of the activity teaching a lesson with an explicit academic focus. (2) *Tutoring*—Tutors or teachers *directly* work with students individually and/or in small groups in order to facilitate the acquisition of skills and knowledge related to concepts addressed during the school day.

surveys returned per program ranged from 16 to 58. One elementary center (Fillmore) and one secondary center (Tyler) were especially apt not to return a complete set of activity leader surveys.

We averaged responses provided by activity leaders in order to explore which types of activities were more prevalent than others across the 11 centers enrolled in the study. As shown in Figure 6, there were some similarities across programs in each grade level category.

Figure 6. For both elementary and secondary programs, working *alone*, working in *small groups*, and exploring *new things* were the *most common* things youth did during activities.



Note: Based on 342 activity leader surveys collected from the 11 study centers during the 2016–17 school year

There were also some notable differences between elementary and secondary centers. In particular, activity leaders in secondary centers were more apt to report that *Youth worked on a group project that will take multiple sessions to complete* (Multiple Sessions); *Youth learned or practiced a skill that is not related to a specific school-day content area* (Non-academic Skill); *Youth participated in a competition, contest, or game* (Competition); and *Youth participated in activities that allowed them to explore and discover new things on their own* (New Things). In addition, although less commonly offered in both elementary and secondary programs, activity leaders in secondary programs were also more apt to report that *Youth delivered a presentation to the whole group or an external audience* (Give Presentation) and *Youth planned future activities or projects* (Planning) than their counterparts working in the elementary sites.

The activities reported to be more prevalent in secondary centers generally seem to represent the greater degree of autonomy and opportunities for agency that can be afforded to early adolescents given the emergence of new cognitive tools and higher order reasoning skills that provide youth an opportunity to play a larger role in pursuing specific learning goals (Larson & Angus, 2011). In this sense, the greater availability of these types of activities would seem to be consistent with the emerging developmental capacities of participating youth. This is a theme

that will emerge again when examining the relationship between observed program quality and youth-reported engagement in programming.

Youth Development Outcomes

Pre-Post Outcomes

The primary goal of the Washington Quality to Youth Outcomes Study was to explore the relationship between observed program quality, reported youth engagement in programming, and changes in a series of youth development-related outcomes. Having discussed data collection activities related to program quality and youth engagement in the preceding sections, we now turn our attention to explaining our approach to youth outcome measurement.

In order to assess the manner in which youth may have changed during their enrollment in 21st CCLC-funded programs in the 2016–17 school year, the *Youth Motivation, Engagement, and Beliefs Survey* was used on a pre-post basis to collect outcome data on youth enrolled in the study. Originally developed by the Youth Development Executive of King County (YDEKC) to support the Road Map Project, a collective impact initiative aimed at increasing educational attainment in the King County, Washington, region, the *Youth Motivation, Engagement, and Beliefs Survey* has been modified by AIR and used in assessing youth functioning on a series of youth development-related outcomes as part of its statewide 21st CCLC evaluation work with OSPI since 2014. A full validation report pertaining to the measure has been previously published by AIR (Naftzger, 2016).

Each administration of the *Youth Motivation, Engagement, and Beliefs Survey* was designed to gauge how well youth were doing in four key areas at the time the survey was taken:

(a) academic identity, (b) positive mindsets, (c) self-management, and (d) interpersonal skills. Items appearing on each of these scales are outlined as follows:

Academic Identity (6 items)

1. Doing well in school is an important part of who I am
2. Getting good grades is one of my main goals
3. I take pride in doing my best in school
4. I am a hard worker when it comes to my schoolwork
5. It is important to me to learn as much as I can
6. I like doing challenging work at school because I know I will learn more

Positive Mindsets (7 items)

1. I finish whatever I begin
2. I stay positive when things don't go the way I want
3. I don't give up easily
4. I try things even if I might fail
5. I can solve difficult problems if I try hard enough
6. I can do a good job if I try hard enough
7. I stay focused on my work even when it's boring

Self-Management (7 items)

1. I can stop myself from doing something I know I shouldn't do
2. When I'm sad, I do something that will make me feel better
3. I can control my temper
4. I can handle stress
5. I can calm myself down when I'm excited or upset
6. When my solution to a problem is not working, I try to find a new solution
7. I think of my past choices when making new decisions

Interpersonal Skills (6 items)

1. I listen to other people's ideas
2. I work well with others on group projects
3. I feel bad when someone gets their feelings hurt
4. I respect what other people think, even if I disagree
5. I try to help when I see someone having a problem
6. When I make a decision, I think about how it will affect other people

Youth taking the survey selected from the following set of response options for each survey item: (1) *Not at all true*, (2) *Somewhat true*, (3) *Mostly true*, or (4) *Completely true*.

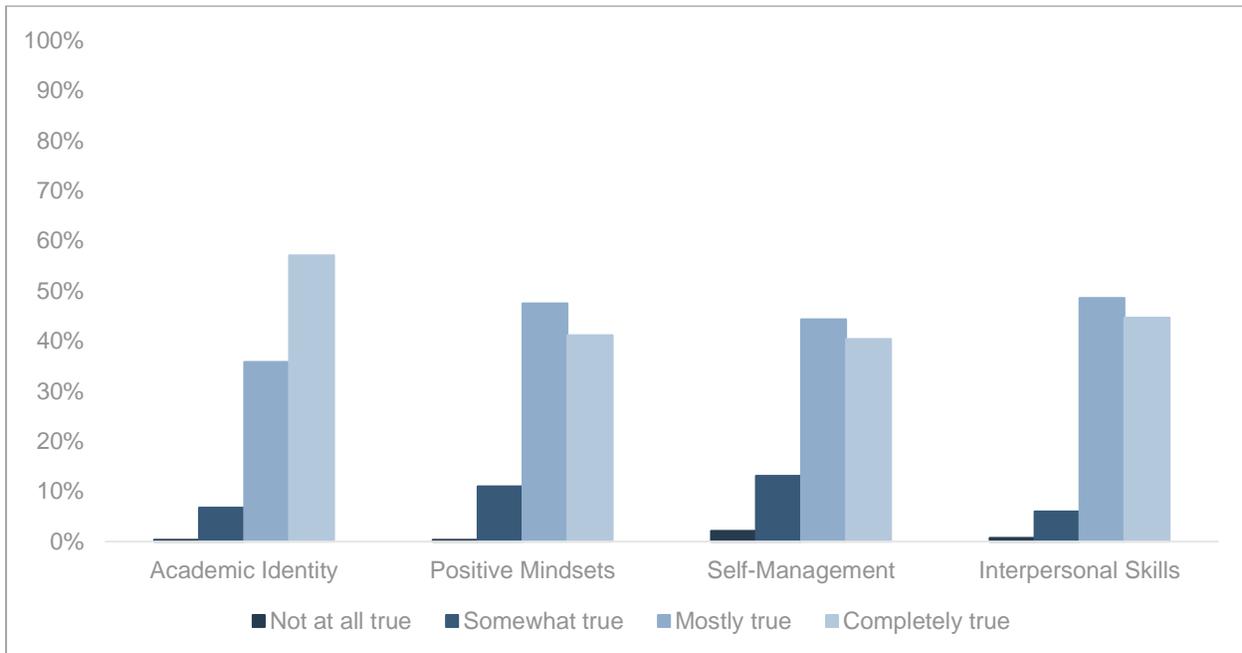
A total of 346 youth in Grades 4–9 from the 11 21st CCLC-funded sites took the presurvey between early October and mid-December 2016, while postsurvey data were collected from 282 of these same youth between late April and early June 2017. On average, a total of 200 days elapsed between presurvey and postsurvey data collection. Youth represented in the sample attended 21st CCLC programming an average of 66 days during this period.

We calculated survey scale scores using the Rasch rating scale model in Winsteps. This approach allowed us to classify youth as falling within one of the rating scale options (i.e., *Not at all true*, *Somewhat true*, *Mostly true*, or *Completely true*) based on their presurvey scale score for a given construct. We considered youth with a score that placed them in the *Completely true* category to be functioning at a relatively high level on the construct in question, while youth falling in the *Not at all true* category were functioning at the lowest level at the time of presurvey data collection.

In order to assess potential growth between presurvey and postsurvey administration, we included youth falling within only the *Not at all true*, *Somewhat true*, and *Mostly true* portions of the rating scale on the presurvey in these analyses. We considered this population of youth to have room to grow because they were not functioning at the top of the rating scale at baseline.

As show in Figure 7, the vast majority of youth had a prescore that put them in the either the *Mostly true* or *Completely true* portion of the survey.

Figure 7. More than 80% of the 282 youth with presurvey and postsurvey data had a baseline score that put them in either the *Mostly true* or *Completely true* range for each scale.



Note: Based on 282 youth outcome surveys collected from the 11 study centers in fall 2016

This meant that the number of youth included in analyses to assess growth on the *Youth Motivation, Engagement, and Beliefs Survey* varied from 121 youth in the case of *Academic Identity* to 168 in the case of *Self-Management*. In this sense, approximately 40% to 57% of youth were already scoring in the highest rating scale category, *Completely true*, at baseline and, therefore, were excluded from analyses to assess the potential growth in functioning on the survey constructs in question. From the perspective of being able to detect growth on survey scales between administrations of the presurvey and the postsurvey, this was less than ideal, raising concerns about our ability to meaningfully detect growth on the scales. This caveat should be kept in mind when reviewing study finding focusing on youth improvement on the survey outcomes considered.

As shown in Table 1, for most of the centers enrolled in the study, anywhere from half to two thirds of pre-post survey respondents fell within *Not at all true*, *Somewhat true*, and *Mostly true* portions of the rating scale on the presurvey, with this percentage being slightly lower for youth attending elementary centers as compared to youth participating in the secondary programs. In addition, the secondary programs also had a larger number of youth taking both the presurvey and the postsurvey, which means that models oriented at assessing growth on each of the survey scales are going to be based on youth attending secondary programs to a greater extent.

Table 1. Percentage of Pre-Post Survey Respondents With Room to Improve by Survey Scale

Center	N	Academic Identity	Positive Mindsets	Self-Management	Interpersonal Skills
Elementary					
Polk	9	44.4%	66.7%	66.7%	44.4%
Harrison	12	50.0%	58.3%	50.0%	50.0%
Fillmore	31	38.7%	51.6%	45.2%	35.5%
Jefferson	16	31.3%	25.0%	62.5%	43.8%
Jackson	24	29.2%	45.8%	54.2%	33.3%
Pierce	11	36.4%	36.4%	27.3%	45.5%
Secondary					
Adams	19	63.2%	84.2%	78.9%	84.2%
Van Buren	21	61.9%	66.7%	61.9%	76.2%
Monroe	51	51.0%	66.7%	66.7%	68.6%
Madison	46	43.5%	60.9%	60.9%	58.7%
Tyler	42	28.6%	61.9%	61.9%	50.0%

Note: Based on 282 youth outcome surveys collected from the 11 study centers in fall 2016

Youth-Reported Program Impacts

In addition to assessing growth on the youth development-related survey scales represented on the *Youth Motivation, Engagement, and Beliefs Survey*, we asked youth taking the postsurvey to select the top three ways they felt they had been impacted through their participation in the 21st CCLC-funded afterschool program. As shown in Table 2, three areas in particular seemed to float to the top in terms of perceived impacts: (1) interest development (e.g., *Discover things I want to learn more about, Find out what I like to do*, etc.), (2) development of a more positive self-concept (*Feel good about myself*), and (3) support the creation of new friendships (*To make new friends*). Youth in both elementary and secondary programs rated interest development highly, while development of a positive self-concept was more often identified by elementary youth. Secondary youth were more apt to identify new friendships as a key impact. These results are quite noteworthy because interest development as a key outcome of afterschool and out-of-school time programs has not traditionally been examined carefully as an outcome of these programs.

Table 2. Percentage of Postsurvey Respondents Indicating a Particular Program Impact

How has this program helped you specifically?	Elementary (n=102)	Secondary (n=172)	Overall (n=274)
Discover things I want to learn more about	47.1%	39.0%	42.0%
Feel good about myself	43.1%	29.7%	34.7%
Find out what I like to do	29.4%	26.7%	27.7%
Think about what I might like to do when I get older	23.5%	27.9%	26.3%
To make new friends	14.7%	30.8%	24.8%
Think about the kinds of classes I want to take in the future	14.7%	29.7%	24.1%
Learn things that will be important for my future	17.6%	27.3%	23.7%
Feel good because I was helping my community	17.6%	16.9%	17.2%
With my confidence	17.6%	16.9%	17.2%
Learn things that will help me in school	18.6%	16.3%	17.2%
Find out what I'm good at doing	15.7%	14.5%	15.0%
Learn about things that are important to my community	18.6%	7.0%	11.3%
Find out what is important to me	9.8%	11.6%	10.9%

Note: Based on 274 youth outcome surveys collected from the 11 study centers in spring 2017

Youth Summative Experiences in Programming

We also added a series of scales to the postsurvey that asked youth to reflect on their afterschool programming experiences in three areas: (1) their perceptions of other youth in the program, (2) their perceptions of the adult activity leaders working in the program, and (3) how frequently they were able to participate in different type of activities that would have afforded them the opportunity to experience a sense of agency. Unlike the youth experience surveys collected throughout the 2016–17 school year, the postsurvey items asked youth to make summative judgments about what their overall afterschool programming experience had been like during the course of the school year.

Almost any quality framework related to afterschool service provision contains criteria related to supporting positive interaction among youth attending programming and with the adults leading program activities. Following are the questions asked in each of these areas to youth taking the postsurvey in spring 2017:

At this program, how do kids get along? Indicate how true each statement is based on your own experience in this program.

- Kids here are friendly with each other.
- Kids here treat each other with respect.
- Kids here listen to what the teachers tell them to do.
- Kids here don't tease or bully others.
- Kids here support and help one another.

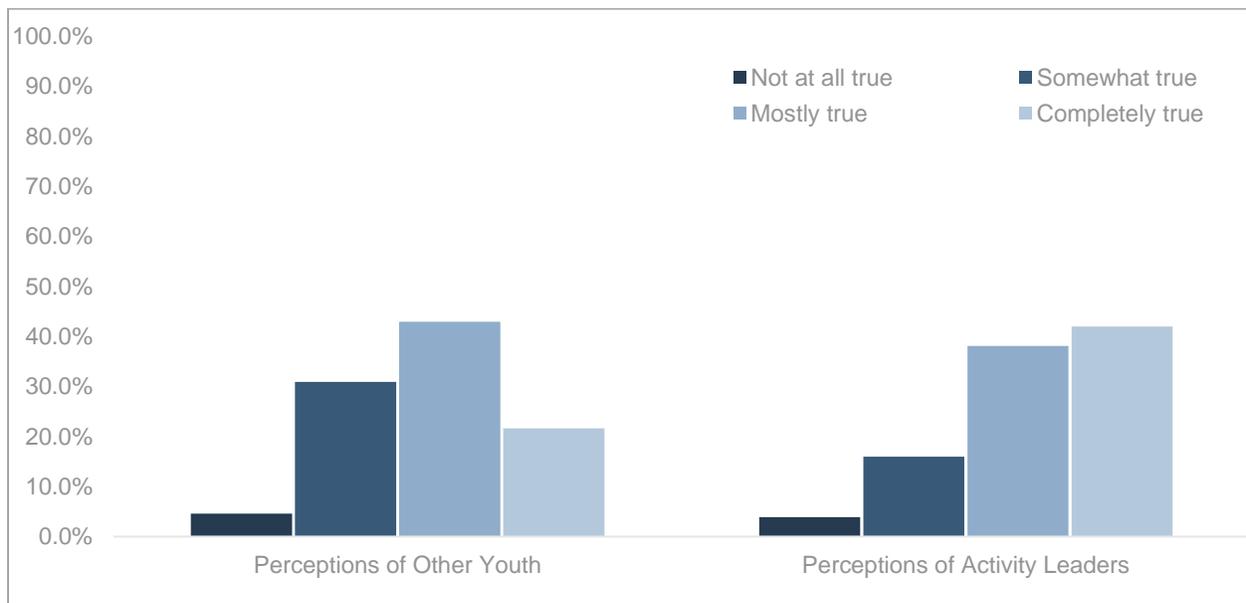
Thinking about the adults in this program, how true are these statements for you? In this program, there is an adult here...

- Who is interested in what I think about things.
- Who I can talk to when I am upset.
- Who helps me when I have a problem.
- Who I enjoy being around.
- Who has helped me find a special interest or talent (something I’m good at).
- Who asks me about my life and goals.
- Who I will miss when the program is over.

For each of these scales, youth responded to each of these items using a 4-point rating scale: (1) *Not at all true*, (2) *Somewhat true*, (3) *Mostly true*, or (4) *Completely true*. Again, we calculated survey scale scores using the Rasch rating scale model in Winsteps. This approach allowed us to classify youth as falling within one of the rating scale options (i.e., *Not at all true*, *Somewhat true*, *Mostly true*, or *Completely true*) based on their postsurvey scale score for each construct. We considered youth with a score that placed them in the *Completely true* category to perceive afterschool activity leaders and other youth in a very positive light, while youth with a score in the *Not at all* range of the scale largely had a negative view of these groups.

In this sense, youth were largely very positive about the activity leaders working in their program. In terms of other youth in the program, while the plurality of respondents fell within the *Mostly true* portion of the scale (43%), another 31% fell in the *Somewhat true* portion of the scale, indicating a more mixed view of other youth in the program.

Figure 8. More than 80% of youth fell within the *Mostly true* or *Completely true* range of the scale for perceptions of activity leaders, while there was more variation in how youth viewed their peers in the program.



Note: Based on 282 youth outcome surveys collected from the 11 study centers in spring 2017

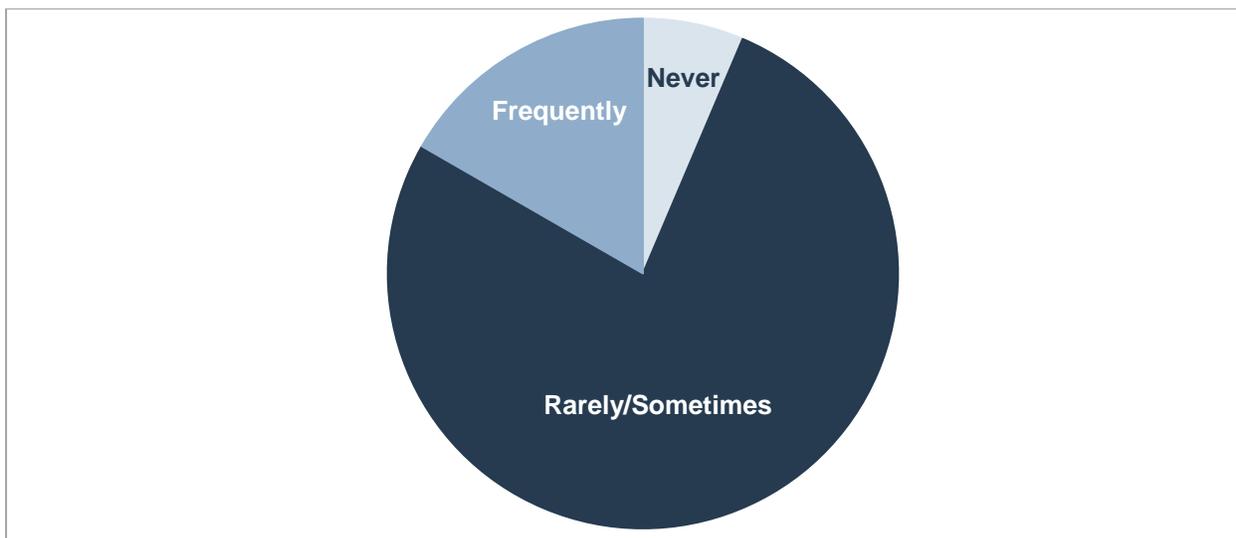
In addition to questions meant to elicit how responding youth felt about the afterschool activity leaders and other youth attending the program, we added a scale to the postsurvey that was designed to assess how frequently youth had the opportunity to participate in activities that allowed them to experience a sense of agency while participating in the program. As noted previously, as youth enter early adolescence, there is substantial growth in the cognitive capacities of youth that allow them to engage in tasks that require higher order thinking and assume a more autonomous role in their learning and in the creation and execution of plans to achieve goals and objectives associated with afterschool activities (Larson & Angus 2011). Our hypothesis going into the study was that more frequent exposure to these types of opportunities would be associated with higher levels of engagement in programming and, as result, greater improvement on the domain of youth development outcomes under consideration. We outline the items asked on this scale below:

Now think about this program in particular. When you are at this program, how often...

- Do you get to choose how you spend your time?
- Can you suggest your own ideas for new activities?
- Do you get to choose which activities you do?
- Do you get to help plan activities for the program?
- Do you get the chance to lead an activity?
- Do you get to be in charge of doing something to help the program?
- Do you get to help make decisions or rules for the program?

In answering these questions, youth responded to each of these items using a 4-point rating scale: (1) *Never*, (2) *Rarely*, (3) *Sometimes*, or (4) *Often*. Calibration of scale scores using the Rasch rating scale demonstrated that youth had a hard time distinguishing between *Rarely* and *Sometimes* when responding to this set of items, so these categories were collapsed into one, as shown in Figure 9. As shown in Figure 9, these types of opportunities were generally not commonly afforded to participating youth.

Figure 9. The vast majority of youth provided responses that put them in the *Rarely/Sometimes* portion of the scale.



Note: Based on 282 youth outcome surveys collected from the 11 study centers in spring 2017

Generally, when considering average center scores across each of the three summative youth perception scales constructed (e.g., perceptions of activity leaders and youth and opportunities for agency), we did not find a tremendous degree of variation from one center to another, although secondary centers had a tendency to demonstrate slightly higher scores on the *Opportunities for Agency* scale. Of some interest was the finding that there was more variation in scores *within* a given center than across centers. This suggests the importance of assessing individual subjective experiences in programming when attempting to understand how these experiences may be related to youth perceptions of their functioning on the youth development outcomes of interest.

Youth Characteristics

Youth enrolled in the study largely mirrored the larger population of youth served by the Washington 21st CCLC program, with the majority of youth eligible for free and reduced-price lunches, scoring below proficiency on either the reading or mathematics portion of the state’s standardized assessment, and were ethnic minorities (see Table 3). Please note that we provide demographic data only for the 282 youth that completed a presurvey and a postsurvey. For most of the analyses described in the following sections, this group will represent the main population of interest when addressing the bulk of the research questions being addressed by the study.

Table 3. Demographic Characteristics Associated With Youth Enrolled in the Study

Youth Characteristic	Percentage
Grade Level	
4th grade	19.2%
5th grade	17.8%
6th grade	22.5%
7th grade	20.7%
8th grade	13.8%
9th grade	6.2%
Gender	
Female	43.0%
Male	57.0%
Other Characteristics	
Ethnic Minority	68.8%
Eligible for Free/Reduced-Price Lunches	70.2%
Limited English Proficiency	23.5%
Enrolled in Special Education Supports	12.5%
Below Proficient in Reading or Mathematics	70.6%
High School-Day Absences*	11.0%

*Defined as missing more than 10% of school days during the previous school year.

Note: Based on 282 youth with pre-post youth outcome survey data available.

Key Study Findings

In order to answer the research questions underpinning the study, the research team ran a series of hierarchical linear models (HLM) to assess the relationship among variables related to program quality, afterschool activities, youth experiences in programming, youth characteristics, and youth outcomes. We ran two general types of models: (1) models where a measure of youth *engagement* derived from the end-of-day youth experiences survey was used as the outcome of interest and (2) models where a pre-post survey-derived, youth development outcome was used as the outcome. We describe findings from running each of these two categories of models in the sections that follow.

In addition, data collection activities undertaken during the course of the study yielded a relatively large number of variables that could have been included in the models designed to answer the study's research questions. To ensure the parsimony of the final models, we conducted a model-building exercise to test the efficacy of various variables being related to the outcomes of interest. Through this process, we dropped some variables because they were not related to the outcomes in question. In going through this process, we retained a series of youth-level demographic and participation variables, and these variables appear in the majority of the models described in the sections that follow:

- Youth participation in 21st CCLC programming for 60 days or more
- Youth eligible for free and reduced-price lunches
- Youth with limited English proficiency
- Youth with a high percentage of school-day absences

It is also important to note that most of the models run were characterized by a relatively small number of programs (in this case centers, $n=11$) and youth (typically youth, $n=121$ to $n=168$). As a result, most of the models run did not meet HLM-related criteria for the use of robust standard errors when examining model results. As a result, we would encourage the reader to consider these results as exploratory.

Process Quality and Summative Youth Experiences Related to End-of-Day Engagement

One of the key things we wanted to learn in conducting the Washington Quality to Youth Outcomes Study was how the level of observed process quality as measured by the YPQA was related to the average level of youth-reported engagement in programming. More specifically, we were interested in answering the following research question: *Is there a relationship between youth-reported daily experiences in afterschool programming and observed levels of program quality?*

For this set of analyses, the outcome of interest was the mean engagement score for those youth with pre-post youth outcome data from the *Youth Motivation, Engagement, and Beliefs* survey. We limited the analysis to this set of youth because this was the population of youth enrolled in the study where parent consent was received that allowed the research team access to school-related records held by OSPI. This allowed us to include key youth demographic information in the model (i.e., high number of school-day absences, eligibility for free/reduced-price lunches, and limited English proficiency). In addition, we limited the analysis to those youth that

completed a minimum of three end-of-day youth experience surveys during the span of the 2016–17 school year. This resulted in a final sample of 162 youth who completed anywhere from three to twelve youth experience surveys during the course of the study period.

In addition to the youth demographic variables included in the model, we included the presurvey score for the *Academic Identity* scale. Our sense is that the collection of data using the *Youth Motivation, Engagement, and Beliefs Survey* offered unique information regarding how youth perceive themselves and therefore warranted inclusion in the model. We found the scales appearing on the survey to be strongly correlated (Pearson correlation ranging from .62 to .74), so we decided to include only one scale from the survey in the model. Given 21st CCLC’s emphasis on supporting academic achievement-related outcomes, we opted to include the *Academic Identity* scale.

Finally, we also included two scales from the youth postsurvey pertaining to youth perceptions of other youth in the program (*Perceptions of Other Youth*) and the extent to which they had participated in activities affording opportunities to experience a sense of agency (*Youth-Reported Opportunities for Youth Agency*). We found a fairly high correlation between this latter scale and youth perceptions of the program’s activity leaders (Pearson correlation = .58), so we included only one of these scales in the model.

As shown in Table 4, we found only two of the predictors to be significantly related to the average level of youth-reported engagement in programming: (1) the mean YPQA score and (2) the *Youth-Reported Opportunities for Agency* score from the postsurvey. Both of these variables were positively related to the average level of youth engagement, meaning the higher the scores on these variables, the more youth reported being engaged in programming. These results are consistent with our hypothesis that **higher levels of process quality and more opportunities for youth to experience a sense of agency are associated with a higher degree of engagement in programming.**

We found only one demographic variable to be associated with the mean level of youth engagement. In this case, **youth eligible for free/reduced-price lunches were more engaged in programming than their peers not eligible for free/reduced-price lunches.** This was a moderately significant relationship. This is a promising result given that one of the main goals of the 21st CCLC program is to extend learning opportunities to youth from high-poverty communities who may not otherwise have access to this type of afterschool programming.

Table 4. Demographic and Process Quality Predictors of Average Youth Engagement

Predictors	Coefficient	Standard Error	p-Value
Program-Level Predictor			
Mean YPQA Score	0.204	0.084	0.038*
Youth-Level Predictors			
Presurvey Academic Identity Score	0.082	0.089	0.377
High Number of School-Day Absences	0.067	0.121	0.591
Eligible for Free/Reduced-Price Lunches	0.197	0.091	0.057+
Limited English Proficiency	-0.204	0.155	0.219
Youth-Reported Opportunities for Agency	0.285	0.061	0.001**
Perceptions of Other Youth	0.059	0.075	0.450

Note: +statistically significant at .10; *statistically significant at 0.05, **statistically significant at .01

To further explore what practices described in the YPQA may be especially associated with higher levels of reported youth engagement in programming, we reran the model detailed in Table 4 individually replacing the Mean YPQA Score with each of the four subscales created when calibrating measure scores: (1) *Positive Emotional Climate*, (2) *Learning Supports and Format*, (3) *Collaboration*, and (4) *Observed Opportunities for Youth Agency*. The significance of the youth-level predictors in these models followed the same pattern outlined in Table 4 (as a result, these results are not shown); however, as shown in Table 5, the only YPQA subscale related to engagement was the *Observed Opportunities for Youth Agency* scale. **In this sense, the Observed Opportunities for Youth Agency scale in particular seems to be driving the positive relationship between the mean YPQA score and mean engagement detailed in Table 4.**

Table 5. YPQA Subscales Used as Predictors of Average Youth Engagement

YPQA Center-Level Predictors	Coefficient	Standard Error	p-Value
Positive Emotional Climate	-0.005	0.060	0.935
Learning Supports and Format	-0.010	0.082	0.908
Collaboration	0.056	0.038	0.178
Observed Opportunities for Youth Agency	0.219	0.074	0.017*

Note: *statistically significant at 0.05

As noted previously in the report, we found a number of key differences between the domain of elementary and secondary centers included in the study, including the mean scores on the *Collaboration* and *Observed Opportunities for Youth Agency* subscales of the YPQA (see Figure 2) and the reported level of youth engagement in programming (see Figure 4). In light of this, we wanted to explore how the results outlined in Tables 4 and 5 would be impacted if we added the elementary status of a program to the model. Because the *Observed Opportunities for Youth Agency* subscale seemed to be especially related to the average level of youth-reported engagement, we opted to add scores from this subscale to the model as a center-level predictor.

As shown in Table 6, the level of average youth-reported engagement in programming was significantly lower in elementary programs relative to the secondary sites enrolled in the study. In addition, while scores on the *Youth-Reported Opportunities for Agency* scale derived from the youth postsurvey remained positively and significantly related to the level youth engagement, the *Observed Opportunities for Youth Agency* subscale from the YPQA became nonsignificant in this model. Our sense here is that there may be other important characteristics of elementary programs that are depressing engagement in programming relative to secondary programs. This includes the club-based model adopted by each secondary site, where youth had the opportunity to pick and attend the activities in which they were most interested. This level of choice was less available to elementary youth, who largely attended whatever activity was being offered on a given programming day.

Table 6. Demographic and Process Quality Predictors of Average Youth Engagement With Grade-Level Status of the Center

Predictors	Coefficient	Standard Error	p-Value
Program-Level Predictor			
Elementary Center	-0.260	0.080	0.013*
Observed Opportunities for Youth Agency	0.064	0.072	0.405
Youth-Level Predictors			
Presurvey Academic Identity Score	0.077	0.094	0.431
High Number of School-Day Absences	0.096	0.117	0.433
Eligible for Free/Reduced-Price Lunches	0.224	0.089	0.031*
Limited English Proficiency	-0.144	0.143	0.340
Youth-Reported Opportunities for Youth Agency	0.251	0.063	0.003**
Perceptions of Other Youth	0.070	0.078	0.393

Note: *statistically significant at 0.05, **statistically significant at .01

Overall, the results outlined in Tables 4 through 6 suggest that the adoption of practices that afford youth opportunities to experience a sense of agency were positively associated with youth-reported engagement in programming. The robustness of this finding is supported by the fact that we found this relationship when employing both observation-based and youth-reported measures of agency, although the finding was most consistent in the case of the latter measure, which suggests the relative importance of understanding how these practices are experienced from the point-of-view of participating youth.

Daily Activities Related to Youth Engagement

We also were interested in examining the relationship between activities reported on by activity leaders taking place on the days youth experience data were collected and levels of youth-reported engagement in programming. In this case, we were interested in answering the following question: *Is there a relationship between youth-reported daily experiences in afterschool programming and the types of learning opportunities undertaken on a given day?*

For this analysis, individual youth engagement scores were the outcome of interest, and we used mean scores for each item appearing on the daily activities survey completed by activity leaders as potential predictors of youth engagement (this was level one in the model). Individual responses provided on a given programming day were nested within youth (level 2), and youth were nested within programs (level 3). To maximize the amount of youth engagement data included in these models, no youth-level demographic variables were included in this model. For this analysis, we nested 1,197 youth engagement scores associated with 97 afterschool programming days where complete daily activities survey data were available within 363 youth providing youth engagement data.

We found two types of activities to be significantly related to youth engagement scores in a positive fashion and therefore we included them in the final model (see Table 7):

- Youth participated in whole-group discussions you facilitated.
- Youth listened to a presentation from a speaker or special guest from outside the program.

Table 7. Daily Activities Predicting Youth Engagement in Programming

Daily Activity	Coefficient	Standard Error	p-Value
Group Discussion	0.148	0.053	0.020*
Listen to a Speaker/Special Guest	0.221	0.091	0.036*

Note: *statistically significant at 0.05

In this sense, the greater time dedicated to these activities, the higher the level of engagement reported by youth when taking the end-of-day youth experience survey. We hypothesize that interest development potentially plays a key role in why the other activities identified in Table 7 are strongly associated with youth engagement in programming, particularly in relation to the secondary sites where youth primarily chose what activities they were going to participate in given the club-based format of these programs. Opportunities for learning to occur through social processes (group discussions) (Hidi & Renninger, 2006) and novelty (listening to an outside speaker) have been linked to interest development in other studies as well (Durik & Harackiewicz, 2007).

Changes in Youth Development Outcomes

To assess the manner in which youth may have changed during their enrollment in 21st CCLC-funded programs during the span of the 2016–17 school year, we examined how pre-post scores changed on the youth development-related outcomes measured on the *Youth Motivation, Engagement, and Beliefs Survey*, specifically *Academic Identity, Positive Mindsets, Self-Regulation, and Interpersonal Skills*.

These analyses allowed us to address the following research question: *To what extent do youth demonstrate improvement on youth development outcomes after participating in afterschool programming?*

As mentioned previously, to assess potential growth between presurvey and postsurvey administration, only youth falling within the *Not at all true, Somewhat true, and Mostly true* portions of the rating scale on the presurvey were included in these analyses. This population of youth were considered to have room to grow because they were not functioning at the top of the rating scale at baseline. As shown in Table 8, this meant that the number of youth included in analyses to assess growth on the *Youth Motivation, Engagement, and Beliefs Survey* varied from 121 youth in the case of *Academic Identity* to 168 in the case of *Self-Management*.

As shown in Table 8, for youth that had room to grow on a given survey construct, there was significant growth across each of the four survey constructs between presurvey and postsurvey administrations based on paired sample *t* tests. We hypothesized that youth attending programming more frequently would be more apt to demonstrate growth on the survey constructs. To explore the extent to which this hypothesis may be true, we split youth with room

to grow on a given construct into three groups based on their level of attendance in the program: (1) less than 30 days, (2) 30 to 59 days, and (3) 60 days or more. As shown in Table 8, only youth in the 60-day or more attendance group demonstrated significant improvement in relation to the *Academic Identity*, *Positive Mindsets*, and *Self-Regulation* scales. Not only do these results seem to support our hypothesis but they also provide programs with a sense of what attendance threshold may be needed in order for participating youth to demonstrate significant improvement on this domain of constructs.

We did not find similar results, however, in relation to the *Interpersonal Skills* scale. In fact, youth attending less than 30 days demonstrated moderately significant improvement on this construct, while there was no significant improvement for youth attending 30 days or more. We plan to examine this further in later sections of the report where we explore how youth-reported interactions with their peers in the program may have influenced growth on this particular construct.

Table 8. Changes in Average Pre-Post Youth Survey Scores by Construct and Program Attendance Level

Construct/Days Attended	N	Presurvey Mean	Postsurvey Mean	p-Value
Academic Identity	121	2.69	2.87	.001**
< 30 days	33	2.71	2.83	.283
30 to 59 days	38	2.71	2.82	.189
60 days+	47	2.68	2.93	.007**
Positive Mindsets	166	2.71	2.79	.011*
< 30 days	44	2.74	2.77	.549
30 to 59 days	53	2.74	2.74	.904
60 days+	64	2.66	2.84	.011*
Self-Management	168	2.62	2.71	.008**
< 30 days	45	2.62	2.63	.870
30 to 59 days	52	2.66	2.71	.364
60 days+	67	2.60	2.77	.008**
Interpersonal Skills	156	2.71	2.80	.025*
< 30 days	42	2.67	2.79	.065+
30 to 59 days	56	2.77	2.81	.386
60 days+	55	2.69	2.78	.263

Note: +statistically significant at .10; *statistically significant at 0.05, **statistically significant at .01

Process Quality, Engagement, and Changes in Youth Development Outcomes

One of the major goals of the Washington Quality to Youth Outcomes Study was to explore the relationship among variables related to program quality, youth experiences in programming, youth characteristics, and changes in the youth development outcomes measured through pre-post administration of the *Youth Motivation, Engagement, and Beliefs Survey*. In running these models, there was particular interest in assessing the relationship between YPQA scores, mean

youth engagement scores, and improvement on each of the youth development outcomes measured through pre-post administration of this survey. The goal of these analyses was to answer the following research question: *Is there a relationship between program quality, youth-reported experiences in programming, and changes in key youth development outcomes?*

Our hypothesis is that we would find a positive relationship between YPQA and engagement scores and changes in the youth development outcomes measured through pre-post survey administration. However, from the models we ran, we did not find a positive relationship between mean engagement scores derived from the end-of-day youth experience survey and changes in the youth development outcomes examined. As a consequence, we dropped the mean engagement score derived from the daily youth experience survey from each model because it was not significantly related to any of the outcomes examined, and there was substantial missing data for this variable because not all youth taking the pre-post survey provided at least three youth experience surveys during the 2016–17 school year.

The relationship between YPQA scores and changes in youth development outcomes was found to be a bit more interesting. When steps were taken to assess the potential main effect of YPQA scores on changes in the youth development outcomes under consideration, we did not find any significant results consistent with our hypothesis that higher YPQA scores would be associated with greater youth improvement. However, as shown in Tables 9 through 15, we did find several positive relationships between summative, youth-reported experiences in programming collected via the postsurvey and youth improvement on the outcomes examined, particularly around youth perceptions of their peers in the program. More specifically, the more youth reported having a positive view of their peers in the program, the more likely growth was demonstrated on the outcomes of interest.

In addition, we took steps to explore whether YPQA scores may have moderated the relationship between variables such as youth perceptions of other youth in the program and changes on the youth outcomes under consideration. Here, we did find a number of instances where higher YPQA scores *strengthened* the relationship between variables such as youth perceptions of their peers in the program and improvements in the youth outcomes examined. Tables will be included below summarizing instances where YPQA scores were found to have a significant moderating role between various youth-level predictors and growth on the youth development outcomes examined.

Tables 9 through 15 highlight results from the final multilevel models run to explore the relationship between youth characteristics and youth experiences in programming and changes in the domain of youth development outcomes. We included the overall mean YPQA quality score in each of these models. We also took steps to test whether individual YPQA subscales were associated with changes in any of the pre-post survey outcomes. Only significant results related the YPQA subscales are shown and are limited to instances where cross-level interactions were assessed.

In addition, in light of the results highlighted in Table 8, we added a variable to each model indicating whether youth attended 21st CCLC programming 60 days or more during the 2016–17 school year.

In Table 9, we outline results for the model that explored changes in the *Academic Identity* scale. In addition to the youth-level predictors included in previous models, we added two additional predictors to this model. Given that this outcome pertains to how youth view school and their ability to do well academically, we thought it would be important to add a variable that indicates whether the youth in question were scoring below proficiency on state assessments in reading and/or mathematics (*Below Proficiency—Reading/Mathematics*). We also thought that the *Youth-Reported Opportunities for Agency* scale would be less applicable to this outcome relative to the *Perceptions of Activity Leaders* scale. Because we found these two scales to be highly correlated, we opted to go only with the *Perceptions of Activity Leaders* scale for this model.

Finally, we added the presurvey score to the model detailed in Table 9 (*Presurvey Academic Identity Score*). Taking this step meant that we can interpret the model results as the degree of growth youth exhibited on the *Academic Identity* scale between presurvey and postsurvey administration. We would expect that the prescore would be positively and significantly related to the postscore. We conducted this step with each of the models run for each of the youth development outcomes under consideration in this section of the report.

Most surprisingly, we found a significant negative relationship between the mean YPQA score and pre-post growth on the *Academic Identity* scale. It is possible that this finding is related to the types of academic offerings made available in the study programs, which may not have scored as well on the YPQA but did help boost youths’ self-concept pertaining to school and their ability to succeed academically. Steps were taken to assess the moderating effect that YPQA scores may have between significant youth-level predictors included in the model and improvements on this scale, but no significant results were found here either. Of particular interest was the significant, positive relationship between the *Perceptions of Other Youth* scale derived from the postsurvey and youth growth on the *Academic Identity* scale (see Table 9). **In this sense, the more youth had a positive view of other youth in the program, the greater degree of growth they demonstrated on the *Academic Identity* scale between presurvey and postsurvey administrations.**

Table 9. Demographic, Process Quality, and Youth Experience Predictors of Changes in the Academic Identity Scale

Predictors	Coefficient	Standard Error	p-Value
Center-Level Predictor			
Mean YPQA Score	-0.310	0.121	0.031*
Youth-Level Predictors			
Presurvey Academic Identity Score	0.306	0.152	0.071+
Attended 21st CCLC Programming for 60 days+	0.003	0.101	0.975
High Number of School-Day Absences	-0.317	0.154	0.066+
Eligible for Free/Reduced-Price Lunches	0.011	0.010	0.915
Limited English Proficiency	-0.091	0.118	0.459
Below Proficiency—Reading/Mathematics	0.152	0.149	0.330
Perceptions of Activity Leaders	0.144	0.088	0.131
Perceptions of Other Youth	0.371	0.086	0.002**

Note: +statistically significant at .10; *statistically significant at 0.05, **statistically significant at .01

We show results for the *Positive Mindsets* scale in Table 10. Unlike the model pertaining to changes in the *Academic Identity* scale, we observed no significant relationship between the mean YPQA score and changes in the *Positive Mindsets* scale. However, we found a significant and positive relationship between youth-reported *Perceptions of Other Youth* in the program and an improvement in the *Positive Mindsets* score between presurvey and postsurvey administrations. We found a similar finding in relation to the *Youth-Reported Opportunities for Agency* scale, derived from the youth postsurvey. **In this sense, the more positive youth viewed their peers in the program and the more frequently they reported engaging in practices that provided opportunities to experience agency, the greater the degree of improvement on the *Positive Mindsets* scale presurvey to postsurvey.** In examining the items making up the *Youth-Reported Opportunities for Agency* scale, we would expect that this scale in particular would be related to the *Positive Mindsets* scale given the affordance of opportunities to take control of one’s learning and pursue specific goals within the confines of the program.

In addition, consistent with the results summarized in Table 8 youth attending 21st CCLC programming for 60 days or more also demonstrated a greater degree of improvement on the *Positive Mindsets* scale relative to youth attending programming for less than 60 days, although this relationship was only moderately significant.

Table 10. Demographic, Process Quality, and Youth Experience Predictors of Changes in the Positive Mindsets Scale

Predictors	Coefficient	Standard Error	p-Value
Program-Level Predictor			
Mean YPQA Score	-0.027	0.076	0.731
Youth-Level Predictors			
Presurvey Positive Mindset Score	0.265	0.145	0.097+
Attended 21st CCLC Programming for 60 days+	0.152	0.072	0.059+
High Number of School-Day Absences	0.000	0.094	0.997
Eligible for Free/Reduced-Price Lunches	0.083	0.073	0.286
Limited English Proficiency	-0.100	0.093	0.308
Youth-Reported Opportunities for Agency	0.263	0.071	0.005**
Perceptions of Other Youth	0.213	0.057	0.005**

Note: +statistically significant at .10; *statistically significant at 0.05, **statistically significant at .01

Steps were also taken by the research team to rerun each of these models to assess the way mean YPQA scores may have moderated the relationship between significant youth-level predictors and changes in each youth development scale. Changes in the *Positive Mindsets* scale was one area where various YPQA scales were found to significantly moderate the relationship between select level-one predictors and growth on this scale. For example, **higher Mean YPQA, YPQA Collaboration, and YPQA Observed Opportunities for Youth Agency scores all were found to strengthen the positive relationship between *Perceptions of Other Youth* and improvements on the *Positive Mindsets* scale (see Table 11).** This was especially the case with the *Observed Opportunities for Youth Agency* and *Collaboration* scales. In the sense, the higher the *Observed*

Opportunities for Youth Agency and *Collaboration* scores, the stronger the positive relationship between youth perceptions of their peers and improvements on the *Positive Mindsets* scale.

A similar finding was found in relation to relationship between *Youth-Reported Opportunities for Agency* (from the youth postsurvey) and improvements on the *Positive Mindsets* scale. In this instance, higher scores on the *Learning Supports and Format* scale of the YPQA were found to strengthen this positive relationship. This was a moderately significant finding.

Table 11. Summary of Significant Cross-Level Interactions Between YPQA Scores and Youth-Level Predictors—Changes in the Positive Mindsets Scale

Predictors	Coefficient	Standard Error	p-Value
Perceptions of Other Youth			
x Mean YPQA Score	0.260	0.138	0.093 ⁺
x Collaboration	0.126	0.053	0.041 [*]
x Observed Opportunities for Youth Agency	0.269	0.104	0.029 [*]
Youth-Reported Opportunities for Agency			
x Learning Supports and Format	0.292	0.151	0.084 ⁺

Note: ⁺statistically significant at .10; ^{*}statistically significant at 0.05, ^{**}statistically significant at .01

As shown in Tables 12 and 14, we found similar results in relation to the *Self-Management* and *Interpersonal Skills* scales when exploring whether YPQA scores and youth-level predictors were related to growth on these scales (predictor main effects). We found the *Perceptions of Other Youth* in the program and *Youth-Reported Opportunities for Agency* scales to be positively and significantly related to improvement on each of these scales, although in the case of the *Opportunities for Agency* scale, this relationship was only moderately significant in relation to the *Self-Management* and *Interpersonal Skills* scales.

Consistent with the findings outlined in Table 8, youth attending programming for 60 days or more were found to demonstrate higher growth on the *Self-Management* scale, although this relationship just missed being moderately significant. However, when YPQA scores were assessed as potential moderators, both the *Mean YPQA* score and the *Collaboration* subscale (moderately significant) were found to strengthen the relationship between youth attending programming for 60 days or more and growth on the *Self-Management* scale (see Table 13).

Table 12. Demographic, Process Quality, and Youth Experience Predictors of Changes in the Self-Management Scale

Predictors	Coefficient	Standard Error	p-Value
Center-Level Predictor			
Mean YPQA Score	-0.016	0.076	0.841
Youth-Level Predictors			
Presurvey Self-Management Score	0.391	0.169	0.044*
Attended 21st CCLC Programming for 60 days+	0.130	0.074	0.107
High Number of School-Day Absences	-0.086	0.169	0.620
Eligible for Free/Reduced-Price Lunches	0.031	0.065	0.646
Limited English Proficiency	-0.036	0.069	0.611
Youth-Reported Opportunities for Agency	0.165	0.080	0.065+
Perceptions of Other Youth	0.320	0.076	0.002**

Note: *statistically significant at .10; *statistically significant at 0.05, **statistically significant at .01

Table 13. Summary of Significant Cross-Level Interactions Between YPQA Scores and Youth-Level Predictors—Changes in the Self-Management Scale

Predictors	Coefficient	Standard Error	p-Value
Attended 21st CCLC Programming for 60 days+			
x Mean YPQA Score	0.449	0.165	0.024*
x Collaboration	0.145	0.076	0.089+

Note: *statistically significant at .10; *statistically significant at 0.05, **statistically significant at .01

Finally, as shown in Table 15, we also found that the *YPQA Learning Support and Format* subscale served to strengthen the relationship between *Youth-Reported Opportunities for Youth Agency* and growth on the *Interpersonal Skills* scale of the youth survey (a moderately significant finding).

Table 14. Demographic, Process Quality, and Youth Experience Predictors of Changes in the Interpersonal Skills Scale

Predictors	Coefficient	Standard Error	p-Value
Center-Level Predictor			
Mean YPQA Score	0.051	0.096	0.607
Youth-Level Predictors			
Presurvey Interpersonal Skills Score	0.387	0.158	0.035*
Attended 21st CCLC Programming for 60 days+	-0.000	0.078	0.998
High Number of School-Day Absences	-0.152	0.138	0.296
Eligible for Free/Reduced-Price Lunches	0.133	0.088	0.160
Limited English Proficiency	0.040	0.103	0.704
Youth-Reported Opportunities for Youth Agency	0.162	0.076	0.058+
Perceptions of Other Youth	0.232	0.090	0.029*

Note: *statistically significant at .10; *statistically significant at 0.05

Table 15. Summary of Significant Cross-Level Interactions Between YPQA Scores and Youth-Level Predictors—Changes in the Interpersonal Skills Scale

Predictors	Coefficient	Standard Error	p-Value
Youth-Reported Opportunities for Youth Agency			
x Learning Supports and Format	0.337	0.154	0.055*

Note: *statistically significant at .10; *statistically significant at 0.05

Overall, we learned several things from these analyses:

- Contrary to our hypothesis, mean engagement scores were not found to be significantly related to growth on the pre-post youth development outcomes examined. This may be related to fact that on average we had only about six completed youth experience surveys per youth included in these analyses, and only about a little more than half of youth met the three or more completed survey threshold we established for engagement data to be assessed. In this sense, our engagement data may have been too sparse to be viably included in these models.
- YPQA mean scores also were not found to be directly related to changes in relation to the pre-post youth development outcomes in the manner hypothesized.
- Both the *Youth-Reported Opportunities for Agency* and *Perception of Other Youth* scales from the youth postsurvey were positively and significantly related to improvements on the outcomes under consideration. **These findings seem to suggest that providing youth with opportunities to experience agency and creating a positive social climate may be related to supporting both program engagement and moving the needle on desirable youth development outcomes.**
- Several YPQA scores also were found to *strengthen* the positive relationship between the *Youth-Reported Opportunities for Agency* and *Perception of Other Youth* scales from the youth postsurvey and the youth development outcomes examined. For example, higher scores on the *Learning Supports and Format* subscale were found to strengthen the relationship between the *Youth-Reported Opportunities for Youth Agency* and growth on the *Positive Mindsets* and *Interpersonal Skills* scales. This makes some intuitive sense and is supported by the literature as well. Larson (2006), for example, talks about the importance of providing an appropriate level of scaffolding when providing youth with opportunities to experience agency akin to some of the practices found on the *Learning Supports and Format* subscale.
- Overall then, these results suggest two key conclusions: (1) The individual experiences of youth are important to consider when assessing growth on the types of outcomes represented on the *Youth Motivation, Engagement, and Beliefs* survey, and (2) the types of practices found on the YPQA can serve to provide a programmatic setting that enhances the relationship between individual experiences and growth on the types of youth development outcomes we are interested in seeing through the provision of afterschool programming.

However, there is one important caveat to keep in mind in relation to these findings. Because the summative youth experience scales were collected on the same survey used to obtain postsurvey results on the youth development outcomes in question, there is some possibility that these relationships may be related to the fact that they were derived from the same tool. In the future, it may be advantageous to collect postsurvey outcome data and data on summative youth experiences in programming at different time points using different surveys in order to eliminate the potential conflation in results potentially resulting through data collected through the same instrument. Although this is a potential issue when examining summative youth experiences in programming relative to the postsurvey youth outcomes, this issue is not present when we examined youth engagement as an outcome in Tables 4-6 because these data were collected using a completely different measure and at various points in time during the 2016–17 school year. As a result, we generally consider findings related to using youth engagement as an outcome more robust than those identified in this section of the report related to the youth development outcomes appearing on the *Youth Motivation, Engagement, and Beliefs Survey*.

Group Discussion and Changes in Youth Development Outcomes

Earlier in this report, we described how certain activities on the daily activity survey completed by activity leaders on the same days youth experience data were collected were related to youth-reported engagement in programming: (1) participating in a group discussion and (2) listening to a speaker/special guest. Given that these activities were related to youth-reported engagement, we were interested in exploring whether the mean center-level score summarizing site adoption of these practices may be related to the domain of youth development outcomes obtained from the pre-post administration of the *Youth Motivation, Engagement, and Beliefs Survey* (i.e., *Academic Identity, Positive Mindsets, Self-Management, and Interpersonal Skills*).

In addition, given the relationship between *Youth-Reported Opportunities for Agency* and improvements on several of the youth development outcomes examined in the preceding section, we opted to include three activities in these models that would seem to especially afford youth the opportunity experiences a sense of agency:

- Youth worked on a group project that will take multiple sessions to complete.
- Youth participated in activities that allowed them to explore and discover new things on their own.
- Youth planned future activities or projects.

We ran a separate model for each of the four outcomes in question and included average scores on each of the activities from the daily activities survey as center-level predictors in each model. In one of the four models, we found that the mean score related to offering *group discussions* was positively related to the outcome in question. This was the *only* activity positively related to the youth outcomes under consideration. As shown in Table 16, we found more frequent offering of activities with group discussion was moderately significant and positively associated with an improvement on the *Interpersonal Skills* scale of the youth survey, a finding that makes some intuitive sense given the nature of the outcome in question. **In this sense, the more frequently activity leaders reported group discussions, the more youth demonstrated improvement demonstrated on the *Interpersonal Skills* scale.** Also in Table 16, for this model, we still see

that scores related to *Youth-Reported Opportunities for Agency* and *Perceptions of Other Youth* are positively associated with youth improvement on the *Interpersonal Skills* scale.

Table 16. Demographic, Activity, and Youth Experience Predictors of Changes in the Interpersonal Skills Scale

Predictors	Coefficient	Standard Error	p-Value
Program-Level Predictors			
Mean Group Discussion	0.597	0.270	0.077 ⁺
Mean Listen to a Speaker/Special Guest	-0.556	0.545	0.355
Mean Multiple Session Group Project	-0.181	0.242	0.488
Mean Explore and Discover New Things	0.017	0.157	0.919
Mean Planned Future Activities or Projects	-0.206	0.198	0.345
Youth-Level Predictors			
Presurvey Interpersonal Skills Score	0.380	0.168	0.047 [*]
Attended 21st CCLC Programming for 60 days+	-0.059	0.104	0.582
High Number of School-Day Absences	-0.209	0.145	0.180
Eligible for Free/Reduced-Price Lunches	0.134	0.088	0.156
Limited English Proficiency	0.056	0.107	0.614
Youth-Reported Opportunities for Youth Agency	0.142	0.076	0.088 ⁺
Perceptions of Other Youth	0.245	0.095	0.028 [*]

Note: ⁺statistically significant at .10; ^{*}statistically significant at 0.05

Summary of Key Findings

Outlined below is a quick summary of what we found both in relation to the program and youth-level characteristics that were associated with youth-reported engagement in programming and improvements on the domain of youth development outcomes examined as part of the study.

- Youth Engagement.* We found a number of program-level characteristics and youth experiences associated with afterschool participation related to youth-reported *engagement* in programming. As anticipated, we found mean YPQA scores to be significantly and positively related to mean levels of youth engagement in programming, although it appears this finding was especially driven by the *Observed Opportunities for Youth Agency* subscale constructed from a subset of YPQA items used in the study. However, once we added the elementary status of the center to the model, the mean YPQA score was no longer a significant predictor of mean youth engagement scores. We hypothesize that there may have been systematic features of elementary centers that suppressed youth-reported engagement in programming, particularly the lack of choice around what activities to attend, a component that was part of each of the secondary programs enrolled in the study given the club-based activity model each employed.

In addition, we found scores from the *Youth-Reported Opportunities for Youth Agency* scale collected from the youth postsurvey were also positively and significantly related to youth-reported engagement in programming, even after we added the elementary status of the center to the model. **Overall, the major take-away here was that the adoption of practices that afforded youth opportunities to experience a sense of agency were positively associated with youth-reported engagement in programming.**

Finally, we found that two activities described on the daily activity survey completed by activity leaders were related to youth-reported engagement in programming:

(1) participating in a group discussion and (2) listening to a speaker/special guest. Our sense is that these activities likely especially support interest development among participating youth in the content being addressed through a given activity.

- *Youth Outcomes.* Contrary to our hypothesis, we did *not* find that the mean level of youth-reported engagement was significantly related to improvement in the youth development outcomes assessed via the *Youth Motivation, Engagement, and Beliefs Survey* (i.e., *Academic Identity, Positive Mindsets, Self-Management, and Interpersonal Skills*). The only predictors we found to be consistently and significantly related to this set of outcomes were two scales from the postsurvey regarding youth summative experiences in programming: (1) *Youth Perceptions of Other Youth* and (2) *Youth-Reported Opportunities for Agency*. In this sense, considering the summative experiences of youth appeared important when exploring the extent to which youth had improved on the youth development outcome under consideration.

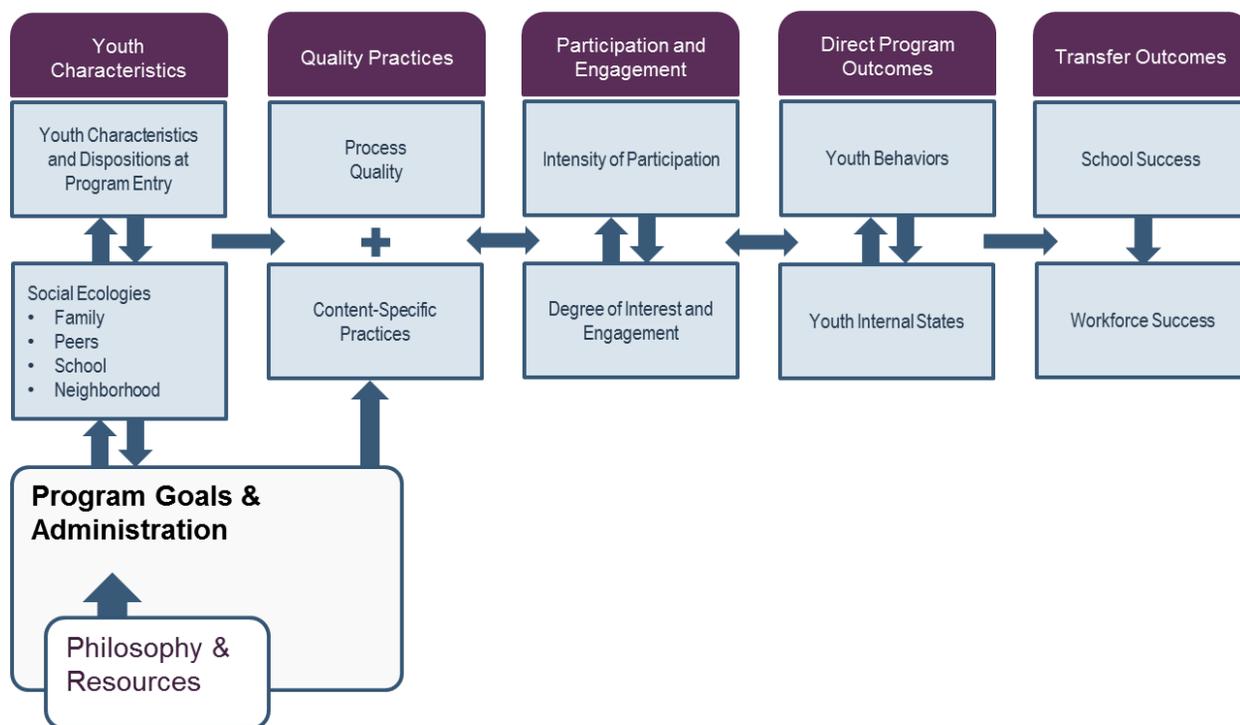
Although we did not find a main effect associated with YPQA scores in terms of growth on the youth development outcomes under consideration, higher YPQA scores were found to significantly strengthen the relationship between the *Youth Perceptions of Other Youth* and *Youth-Reported Opportunities for Agency* scales and growth on the *Positive Mindsets, Self-Management, and Interpersonal Skills* scales. This set of findings demonstrates the potential importance of assessing youth outcomes while also examining the interplay between individual youth experiences in programming and the quality of the programmatic setting.

Finally, the more center activity leaders reported offering activities characterized by *group discussions*, the more youth demonstrated improvement on the *Interpersonal Skills* scale of pre-post youth survey, a finding that makes intuitive sense given the nature of the activity and the outcome in question.

An Alternate Interpretation

In undertaking the analyses outlined in the report, we have tried to follow the conceptual framework introduced at the beginning of the report outlining a process by which youth may benefit from participation in afterschool programs (see Figure 10). In models where mean youth engagement was the outcome of interest (see Tables 4-6 in particular), YPQA scores were used to instantiate *Process Quality* from the conceptual framework. In a similar fashion, scores on the *Youth-Reported Opportunities for Agency* scale derived from responses from the post-youth survey were also treated as a measure of *Process Quality*. We hypothesized that higher scores on each of these measures would be associated with higher levels of youth-reported engagement in programming, a hypothesis that was found to be consistent with our results.

Figure 10. Conceptual Framework



However, it is important to point out that youth engagement data was collected *prior* to post-youth survey data. In this sense, the degree to which youth experienced engagement on a day-to-day basis while participating in programming may have also influenced how they responded to the *Youth-Reported Opportunities for Agency* scale. A similar case can be made that day-to-day experiences in programming also likely influenced how youth responded to the *Perceptions of Other Youth* scale also contained on the post-survey. In this sense, it may be justifiable to treat these post-survey scales as outcomes that could be influenced by youth’s day-to-day engagement in programming. We did this in a series of additional models we ran, using mean youth engagement scores as a predictor, as well as the other demographic-related, youth-level predictors we have included in the models throughout this report.

As shown in Tables 17 and 18, the mean youth engagement score was found so be significantly and positively related to both the *Youth-Reported Opportunities for Agency* and *Perceptions of Other Youth* scales. In this sense, higher engagement scores were associated with higher scores on each of these postsurvey scales.

Table 17. Demographic and Mean Youth Engagement Predictors of the Youth-Reported Opportunities for Agency Scale

Predictors	Coefficient	Standard Error	p-Value
Youth-Level Predictors			
Mean Youth Engagement	0.499	0.090	0.000***
Attended 21st CCLC Programming for 60 days+	-0.088	0.186	0.647
High Number of School-Day Absences	-0.243	0.165	0.173
Eligible for Free/Reduced-Price Lunches	-0.086	0.142	0.555
Limited English Proficiency	0.179	0.123	0.174

Note: ***statistically significant at 0.001

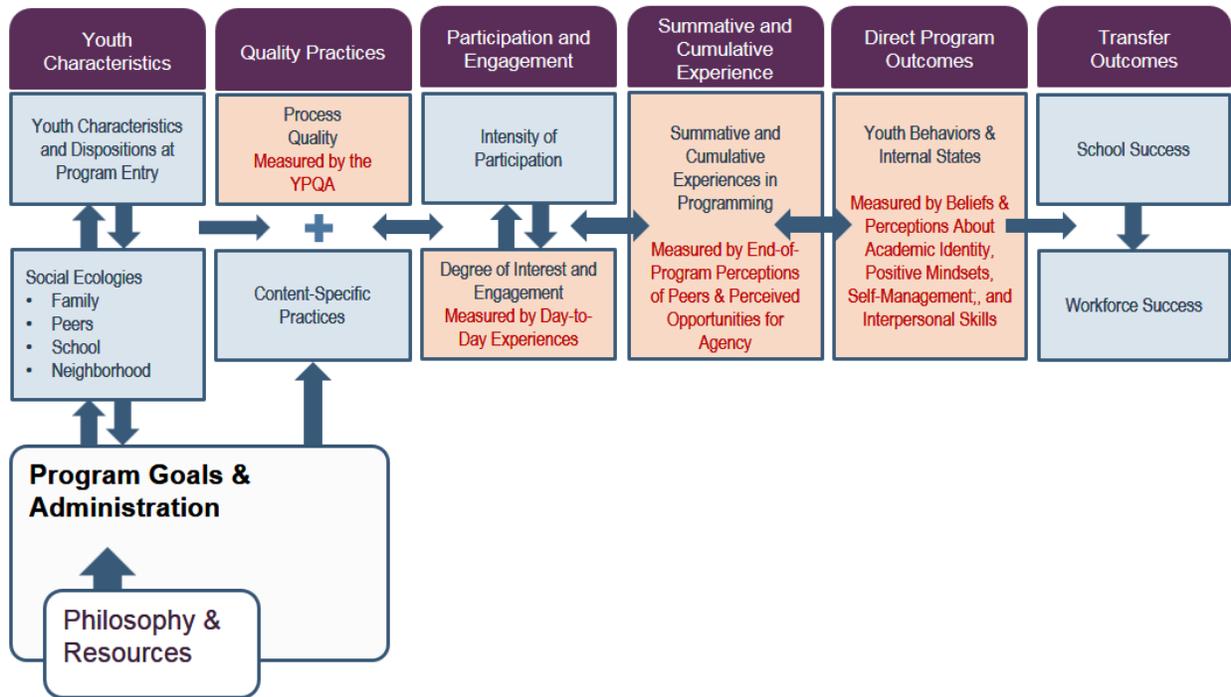
Table 18. Demographic and Mean Youth Engagement Predictors of the Perceptions of Other Youth Scale

Predictors	Coefficient	Standard Error	p-Value
Youth-Level Predictors			
Mean Youth Engagement	0.370	0.118	0.011*
Attended 21st CCLC Programming for 60 days+	-0.209	0.132	0.146
High Number of School-Day Absences	0.016	0.205	0.938
Eligible for Free/Reduced-Price Lunches	-0.014	0.157	0.931
Limited English Proficiency	0.140	0.155	0.389

Note: *statistically significant at 0.05

With these results, it seems like there may be some preliminary evidence that suggests a slightly revised chain of events when trying to understand the relationship between afterschool program participation and youth outcomes as shown in Figure 11. Earlier in this report, we described how higher YPQA scores were associated with higher reported average engagement in programming on a day-to-day basis (see Tables 4-5). Based on the results outlined in Tables 17 and 18, higher mean engagement scores were found to be related to higher scores on both the *Youth-Reported Opportunities for Agency* and *Perceptions of Other Youth* scales from the post-youth survey, suggesting day-to-day engagement in programming is related to youth descriptions of their summative and cumulative experiences in programming. Finally, these summative and cumulative experiences in programming were found to be related to improvements on the domain of youth development outcomes examined in this report (see Tables 9-15). We find this sequence of significant relationships connecting program quality to positive youth experiences in programming to improvement on youth development outcomes to be of particular interest. If anything, this sequence of events may provide an initial template to support future measurement efforts to connect afterschool program quality to improvements on youth development outcomes assessed as part of this study.

Figure 11. Adjusted Conceptual Framework – Focus on Measurement Approaches



Limitations

Although a number of findings derived from the study were significant and consistent with many of our hypotheses regarding the relationship between program quality, youth engagement, and growth on youth development outcomes, there are some important limitations associated with the study that are important to keep in mind.

- *Descriptive study.* The analyses we conducted were descriptive and correlational. In this sense, while we did find evidence of key relationships we expected to find, the reader should not interpret these results as certain practices causing certain outcomes to happen. Our research design does not support this level of inference.
- *Relatively little YPQA data were collected as part of the study.* Programs were observed on one to two days during the 2016–17 school year for the Washington Quality to Youth Outcomes study. As shown by variation in the youth engagement data collected as part of the study, we know programming can vary greatly from one day to the next, both in terms of underlying quality and in terms of the degree to which youth are engaged in programming. It may be the case that we simply had too small a sample of programming days from which to accurately infer overall center-level quality.

- *Relatively small sample sizes.* Our study sample was quite small and based on convenience, particularly at the program level with only 11 sites in total. This certainly may have impeded our ability to detect significant effects, an issue that was compounded by a relatively small pool of youth that completed both the pre-post youth survey and had provided meaningful engagement data through the collection of end-of-day youth experience surveys.
- *High performance on the youth development outcomes measured on the presurvey.* In excess of 80% of youth fell within the top two rating scale categories on the presurvey, indicating a relatively high level of functioning as baseline. Although we identified anywhere from 43% to 60% of youth as having some room to grow on the youth development outcomes measured by the pre-post youth survey, this restricted range may have further limited efforts to detect a potential significant relationship between process quality, mean youth engagement in programming, and growth on the outcomes examined.
- *Concerns about regression to the mean.* Given that youth selected for inclusion in the outcome models had room to improve on the youth development outcomes examined, it may be the case that their improvement was simply representative of regression to the mean as opposed to true improvement on the scale in question.
- *Common method variance.* Scales related to youth-reported, summative experiences in programming (e.g., perceptions of other youth in the program, youth-reported opportunities to experience agency, etc.) were collected using the same survey containing the postsurvey version of the youth development outcomes examined as part of the study. It is possible the significant relationship found to exist between youth experiences and outcomes may be related to the fact that they appeared on the same instrument, thereby inflating the presence of a significant relationship.

When evaluating the results from the study, it is important for readers to keep these caveats and limitations in mind.

Conclusions

The Washington Quality to Youth Outcomes Study focused on trying to understand the relationship between *process quality* instantiated by the YPQA, the degree of youth *engagement* in programming reported by youth, and *youth self-assessments* of how they were functioning on a series of youth development-related outcomes. Our core hypothesis was that higher YPQA scores would be associated with greater youth-reported engagement in programming and that each of these constructs would be associated with improved scores on the four outcome areas measured using a pre-post administration of the *Youth Motivation, Engagement, and Beliefs Survey* (i.e., academic identity, positive mindsets, self-management, and interpersonal skills).

Our initial hypothesis linking process quality to youth-reported engagement in programming was largely supported, with mean YPQA scores significantly and positively related to mean engagement scores. In this sense, higher YPQA scores meant higher youth-reported engagement in programming. In particular, scores on the YPQA subscale that focused on opportunities for youth to experience a sense of *agency* while participating in afterschool activities was especially related to youth-reported engagement in programming and appeared to be the driving force behind why the YPQA was related to mean youth engagement.

We also asked youth on the postsurvey to indicate how frequently they had the opportunity to participate in activities that would afford them the opportunity to experience a sense of agency while participating in programming. Here again, the extent to which youth said they participated in activities that would have afforded them the opportunity to experience a sense of agency was strongly related to mean engagement scores across each of the models created to assess these relationships.

We also hypothesized that the average level of youth-reported engagement in programming and higher program-level YPQA scores would be related to improvement on each of the youth outcomes measured by pre-post scales from the *Youth Motivation, Engagement, and Beliefs Survey*. In terms of a potential link between youth-reported engagement and youth outcomes, the data examined as part of the study did not support this hypothesis. This finding may be related to the relatively sparse nature of our youth-level engagement data.

However, when we asked youth about their overall program experiences on the youth postsurvey, we found youth perceptions of their peers to be positively related to growth on each of the four youth development outcomes examined, suggesting the social climate of the program may be critical to moving the needle on these outcomes. In addition, we found the youth-reported opportunities for agency scale to be positively related to growth on the *Self-Management* and *Interpersonal Skills* scale (moderately significant) and especially on the *Positive Mindsets* scale.

We also did find some evidence linking YPQA scores to changes in the youth development-related outcomes, but not necessarily in the manner initially hypothesized. Although we did not find any main effects linking YPQA scores to changes in the pre-post youth survey outcomes examined, we did find that select *YPQA scores seemed to strengthen* the positive relationship between youth-reported experiences in programming from the youth postsurvey and growth on three of the four youth outcomes (*Positive Mindsets, Self-Management, and Interpersonal Skills*).

This largely occurred in relation to youth perceptions of (a) other youth attending programming in terms of how supportive and friendly their peers in the program were and (b) the opportunities they had to experience a sense of agency while participating in programming. Higher YPQA scores strengthened the relationship between these experiences and growth on the youth development outcomes examined.

These findings also suggest that accounting for individual youth experiences in programming may be especially critical when attempting to assess youth growth on outcomes akin to those assessed in our study and that higher program quality at the setting level can serve to strengthen the connection between youth experiences in programming and positive youth outcomes.

Relatedly, a series of additional analyses demonstrated the presence of a possible chain of events that seemed to better describe the relationship between afterschool program participation and youth outcomes than perhaps what was shown in our initial conceptual framework. Based on results from the study, this new sequence of events looks something akin to the following:

- Higher YPQA scores were associated with higher reported average engagement in programming on a day-to-day basis;
- Higher levels of engagement in day-to-day programming was found to be related to more positive youth descriptions of their summative and cumulative experiences in programming based on post-youth survey results, both in terms of positive peer interactions and more frequent opportunities for agency;
- More positive youth-reported peer interactions and greater opportunities for agency were found to be associated with greater growth on the youth-development outcomes examined.

We find this sequence of significant relationships connecting program quality to positive youth experiences in programming to improvement on youth development outcomes to be of particular interest. If anything, this sequence of events may provide an initial template to support future measurement efforts to connect afterschool program quality to improvements on youth development outcomes assessed as part of this study.

Providing youth with opportunities to experience a sense of agency emerged repeatedly as a significant predictor in models related to both youth engagement and improvement on the post-youth development outcomes, particularly when these experiences were reported directly by youth. We think this is a critical study finding that may warrant further consideration both in how we think about quality tools like the YPQA and in how we try to measure both implementation and the outcomes derived from participation in afterschool programs.

Finally, we were struck by how frequently youth reported impacts related to program involvement centered on the development of new interests. Our sense is this area really has not been investigated fully in terms of how the sparks of interest created in these programs may result in a series of cascading effects relative to the types of opportunities youth seek out in the future and how they may interact differently with other learning environments such as the school

day. We certainly believe there are opportunities to explore this area more fully, particularly in contributing to a more holistic understanding of how afterschool programs may be having a positive impact on youth.

Next Steps

The dataset assembled as part of the Washington Quality to Youth Outcomes Study is relatively unique within the field of afterschool given the focus on measuring both youth day-to-day and cumulative experiences in programming combined with an assessment of growth on a series of youth development-related outcomes. As a result, there are a number of questions beyond the scope of this report that could be further explored through additional analysis of the study dataset. For example, the daily youth experience surveys administered as part of the study measured a number of areas beyond youth engagement in programming, including questions pertaining to the degree to which youth felt (a) challenged by the activities they participated in; (b) they had learned something or got better at something; (c) that what they were doing was relevant to their life and goals; and (d) happy and excited by what they had done in the program on the day in question. Given time and resource constraints, we were not able to fully explore how these experiences were connected to either program quality or changes on the youth development outcomes examined as part of the study. We think exploring these relationships would help to further understand what kinds of afterschool programming experiences may lead to the type of growth we would prefer to see on youth development-related outcomes.

In addition, given 21st CCLC's focus on school-related outcomes, it would be ideal if we could further explore how changes on the youth development outcomes derived from the *Youth Motivation, Engagement, and Beliefs Survey* may be related to academic achievement and school behavior outcomes documented in OSPI's data warehouses. This information would be particularly helpful in understanding what intermediate outcomes may be especially important to cultivate within programs seeking to further promote positive school-related outcomes.

References

- Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology, 99*(3), 597.
- Durlak, J. A., Mahoney, J. L., Bohnert, A. M., & Parente, M. E. (2010). Developing and improving after-school programs to enhance youth's personal growth and adjustment: A special issue of AJCP. *American Journal of Community Psychology, 45*(3-4), 285–293. doi: 10.1007/s10464-010-9298-9
- Durlak, J. A., & Weissberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.
- Kauh, T. (2011). *AfterZone: Outcomes for youth participating in Providence's citywide after-school system*. New York, NY: Public/Private Ventures. Retrieved from <http://www.wallacefoundation.org/knowledge-center/Documents/AfterZone-Outcomes-Youth-Participating-Providences-Citywide-After-School-System.pdf>
- Larson, R. (2006). Positive youth development, willful adolescents, and mentoring. *Journal of Community Psychology, 34*(6), 677–689.
- Larson, R. W., & Angus, R. M. (2011). Adolescents' development of skills for agency in youth programs: Learning to think strategically. *Child Development, 82*(1), 277–294.
- Miller, B. M. (2007). What counts in after school? Findings from the Massachusetts Afterschool Research Study (MARS). *Journal of Youth Development, 1*(3), 98–114.
- Naftzger, N. (2016). *Motivation, Engagement, and Beliefs Survey validation report*. Washington, DC: American Institutes for Research. Retrieved from <https://ydekc.org/wp-content/uploads/2016/11/Final-YDEKC-Survey-Validation-Report.pdf>
- Naftzger, N., Devaney, E., & Foley, K. (2014). *Summary of analyses related to NAZA program outcomes*. Naperville, IL: American Institutes for Research.
- Naftzger, N., Hallberg, K., & Tang, Y. (2014). *Exploring the relationships between afterschool program quality and youth outcomes: Summary of findings from the Palm Beach County quality improvement system*. Naperville, IL: American Institutes for Research.
- Naftzger, N., Nistler, M., Manzeske, D., Swanlund, A., Shields, J., Rapaport, A., Sugar, S. (2013). *Texas 21st Century Community Learning Centers: Year two evaluation report*. Naperville, IL: American Institutes for Research.

Pierce, K. M., Auger, A., & Vandell, D. L. (2013, April). *Narrowing the achievement gap: Consistency and intensity of structured activities during elementary school*. Paper presented at the Society for Research in Child Development Biennial Meeting, Seattle, WA.

Shernoff, D. J., & Vandell, D. L. (2007). Engagement in after-school program activities: Quality of experience from the perspective of participants. *Journal of Youth and Adolescence*, 36(7), 891.

Trill, S., Brohawn, K., & Caruso, C. (2013). *More and better learning: Year One report on ExpandedED Schools*. New York, NY: The After-School Corporation.

ABOUT AMERICAN INSTITUTES FOR RESEARCH

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



AMERICAN INSTITUTES FOR RESEARCH®

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

www.air.org

Making Research Relevant

LOCATIONS

Domestic

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

International

El Salvador
Ethiopia
Haiti
Honduras
Kyrgyzstan
Tajikistan
Zambia