2016 CYFAR Evaluation Report

Submitted by:

The Children, Youth, and Families At-Risk
Professional Development and Technical Assistance (CYFAR PDTA) Team
University of Minnesota & Pennsylvania State University

Lynne M. Borden, Ph.D. (Co-PI)
Daniel F. Perkins, Ph.D. (Co-PI)

Jan Carroll, Ph.D.
Nancy Deringer, Ph.D.
Autumn Guin, M.S.
Amy Gunty, M.A.
Kimberly McCarthy, B.S.
Mark Otto, B.S.
Laura Palmer, M.S.
Shane Potter, M.S.
Matt Roberts, M.S.
Cathy Straub, M.Ed.
Jennifer Wells Marshall, Ph.D.

For additional information, please contact:

Lynne M. Borden, Co-PI
Family Social Sciences
University of Minnesota
lmborden@umn.edu
(612) 624-7707

Developed in collaboration with The United States Department of Agriculture and The National Institute of Food and Agriculture under The University of Minnesota Award No. 2013-05934.

Printed by the University of Minnesota

Table of Contents

EXECUTIVE SUMMARY

In 2017, the Children, Youth, and Families At-Risk Initiative (CYFAR) celebrates its 25th Anniversary. CYFAR provides funding for local Sustainable Community Projects (SCPs) that promote positive outcomes among vulnerable populations throughout the United States and Territories. CYFAR also funds the CYFAR Professional Development and Technical Assistance (CYFAR PDTA) Center, which provides professional development, technical assistance, and evaluation support for CYFAR SCPs. Evaluations of CYFAR SCPs began in 2010 as a pilot program, followed by a mandate that all SCPs collect common measures data starting in 2014. This report was developed to provide an update on the status of the data collection effort, data-based analysis and findings related to outcomes, and recommendations for adaptations to the cross-site evaluation process.

Findings

Data are gathered from SCP participants in four key areas: demographics and program participation level, program quality, core competencies, and short-term outcomes. These participants complete pre-surveys (at the beginning of a program) and post-surveys (at the end of a program), which allows for an analysis of participant change over time while taking part in a CYFAR SCP.

Data Collection. Data collection has been steadily increasing since the pilot program began in 2010, particularly in the years since the collection was mandated in 2014. The number of SCPs submitting data has increased almost every year since 2011, and the number of participants from whom data has been collected has increased from 23 participants in 2011 to 3,350 participants in 2016.

Demographic Differences. There were demographic differences in several outcome measures at the time of the pre-surveys. Female participants scored significantly higher on all core competencies at pre-survey than male participants. Non-Latino or Non-Hispanic participants scored significantly higher on critical thinking at pre-survey than did Latino or Hispanic participants. White youth scored significantly higher than Black youth with regard to caring at pre-survey, and Black youth scored significantly higher than White youth with regard to critical thinking at pre-survey.

There were also demographic differences in scores on common measures of short-term outcomes at the time of the pre-survey. Females scored higher than males on measures of leadership and nutrition. In addition, for middle- and high-school youth, males scored higher than females on the science measure. There were not many differences based on race or ethnicity on common measures of short-term outcomes; however, youth identifying as Non-Hispanic or Non-Latino scored higher on workforce preparation than those identifying as Hispanic or Latino.

The Relationship between Program Quality and Youth Outcome Measures. There was a significant relationship between youth ratings of program quality (post-survey) and all core competencies. Youth who rated program quality as being higher tended to score higher on all core competencies. With regard to short-term outcomes, significant relationships were found between program quality and leadership, high-school nutrition, and middle- and high-school workforce preparation. Overall, when youth rated program quality higher, they were also more likely to report higher scores on these measures of short-term outcomes. Due to the nature of these analyses, however, the causal direction of this relationship is unknown.

Difference in Pre- and Post-Surveys. Scores of core competencies and short-term outcomes were compared between pre- and post-surveys. Results were mixed. Two core competencies (social conscience and personal values) demonstrated a statistically significant, albeit small, decrease between pre-survey and post-survey groups, whereas two other core competencies (decision-making and critical thinking) demonstrated an increase between pre-survey and post-survey groups. Furthermore, several common measures of short-term outcomes demonstrated increases between pre-survey and post-survey groups: leadership development, parenting, high-school workforce preparation, technology for youth in grades three through five, and science for youth in grades three through five.

Limitations. There were some limitations that may affect analyses and conclusions. One major limitation was SCP's infrequent use of the participant-naming convention. This naming convention allows analysts to pair individual participants' pre-surveys with their post-surveys to examine the change for particular participants during their time in an SCP. Since programs often did not follow the naming convention, analysts were unable to match cases; therefore, findings in this report are based on a less robust analysis, comparing the complete group of pre-surveys to the complete group of post-surveys.

Recommendations

Add demographics to post-test surveys. Collecting demographic data at the post-survey will increase the ability to analyze differences in youth outcomes based on demographics. It will also increase the ability to match individual pre-surveys with post-surveys. Furthermore, this allows an evaluation of whether demographic differences on CYFAR common measures are reduced at the end of a program.

Redesign participation level questions. The participation level data obtained for the current report varied greatly among SCPs and participants, which resulted in an inability to use these measures in analyses. It is necessary to adapt participation-level questions in order to analyze these data.

Reduce the current survey size. In order to decrease the amount of time and effort spent on data collection and to increase the quantity and quality of data collected, the surveys should be shortened. Two major changes are recommended: eliminate assessing program quality during the pre-survey, and reduce the number of core competency areas assessed. These changes will simplify the data collection process and provide cleaner data in the future.

Future Directions

The CYFAR PDTA Center has developed and is currently piloting three tools to simplify the data collection process and provide higher quality data.

ERIKA (Electronic Response to Information, Knowledge, and Action): This is an online, illustration-based evaluation tool designed to simplify data collection for youth between the ages of six and ten.

C-NAT (CYFAR Nutritional Assessment Tool: This online health-screening tool assists in the collection of data on health promotion behaviors (e.g., screen time and consumption of fruits and vegetables) for youth between the ages of six and ten.

Evaluation Institute: This consists of a series of instructional modules designed to improve data collection and evaluation capacity among CYFAR grantees.

CYFAR Evaluation Report

In 2017, the Children, Youth, and Families At-Risk Initiative (CYFAR) celebrates its 25th Anniversary. CYFAR provides funding for local Sustainable Community Projects (SCPs) that promote positive outcomes among vulnerable populations throughout the United States and Territories. CYFAR also funds the CYFAR Professional Development and Technical Assistance (CYFAR PDTA) Center, which provides professional development, technical assistance, and evaluation support for CYFAR SCPs. The CYFAR PDTA Center began a pilot evaluation process in 2010 to collect cross-site evaluation data from the community-based programs. The pilot evaluation process lead to a mandate that all programs contribute to this cross-site evaluation design, beginning in 2014. This report presents the findings from this cross-site evaluation initiative.

Methodology

The CYFAR PDTA Center coordinates data collection for CYFAR SCPs. The Center has developed a variety of tools, which are available on the CYFAR website (https://cyfar.org), to assist with the data collection process. Included among these tools is Survey Builder, a program that allows evaluators to use CYFAR common measures to create a questionnaire specifically tailored to each SCP. There are four groups of CYFAR common measures: demographics and participation level (4 measures), program quality (2 measures), core competencies (1 measure), and common measures of short-term outcomes (13 measures). More detailed information about each measure, as well as the instruments themselves can be found at: https://cyfar.org/ilm_common_measures.

Data Requirements

Within the mandate that all SCPs contribute to the CYFAR-wide, cross-site evaluation design, there are certain requirements regarding which measures must be used for different SCPs, based upon the age of the population the SCP serves.

Demographics and Participation Level

These measures are required of all SCPs, regardless of the population served. Demographics are currently only collected as part of the pre-survey and are not collected as part of the post-survey. Participation level is collected as part of both the pre- and post-surveys.

Program Quality

Program quality is required for all SCPs serving youth in middle and high school. The adult version is optional, and is primarily directed toward youth workers in an SCP serving youth. This measure is optional for youth in grades three through five. There is currently no version of this measure for children under third grade.

Core Competencies

Like program quality, measures of core competencies are required for all SCPs serving middle- and high-school youth. The measure is optional for youth in grades three through five. There is currently no version of this measure for children under third grade or for adults.

Common Measures of Short-Term Outcomes

Each SCP is required to choose at least one common measure of short-term outcomes to use in its survey; these measures target areas relevant to each specific SCP. There are six options for high-school youth, five options for middle-school youth, three options for youth in grades three through five, and two options for adults. There are currently no options for measures of short-term outcomes for children under third grade.

Scale Content and Psychometrics

Reliability is key in establishing high-quality measures for use in cross-site evaluation. Reliability is measured by testing the strength of the relationships among items within a scale. For example, if a participant scores high on one item within a scale, they should typically score high on other items within the same scale. The Cronbach's Alpha (α) coefficient was used to estimate the reliability of the CYFAR common measures scales that evaluate program quality, core competencies, and short-term outcomes. Reliability was measured for both the pre-survey and post-survey data on all 32 scales and subscales used within these CYFAR common measures: core competencies (5 scales), program quality (6 scales), and common measures (21 scales).

Demographics and Participation Level

There are four measures in this group: youth demographics, youth participation level, adult demographics, and adult participation level. The demographic measures collect program participant data including gender, age, ethnicity, race, grade level (for youth), parental military service (for youth), education (for adults), employment (for adults), and military service (for adults). The participation level measures collect data regarding frequency and length of participation in the SCP, length of participation in 4-H, and involvement in other activities (including in-school and out-of-school activities for youth). Due to the nature of these measures, there is no estimate of reliability for them.

Program Quality Scale Reliability Estimates

Based upon recommendations from the 2013 CYFAR Evaluation report, the Program Quality Instrument (PQI) was reduced from 10 subscales with a total of 40 items to 5 subscales with a total of 22 items. The current PQI measures several aspects of program quality: physical and psychological safety, supportive relationships, positive social norms, support for efficacy and mattering, and opportunities for skill building. There are two versions of this measure, one for adult youth workers and one for youth participants.

Table 1. Reliability estimates for individual and Total Program Quality scales ranged from moderate to high at both pre-survey and post-survey

The reliability estimates of the PQI subscales ranged from moderate to high at both the pre-survey and post-survey (see Table 1). The Supportive Relationships subscale and the Support for Efficacy and Mattering subscale had the lowest reliability estimates; however, these subscales were composed of three items each. When a scale is composed of only a few items, reliability estimates tend to be lower. For all subscales, pre-survey reliability estimates were similar to the post-survey estimates. Overall program quality (i.e.,

Program Quality [Total]) is an average of all items from the PQI; its reliability estimate was very high at both pre-survey and post-survey.

Table 1. Reliability Estimates for Program Quality Scales					
Scale	**Items**	**Pre-Survey**		**Post-Survey**	
		n	**α**	***n***	**α**
Physical and Psychological Safety	6	1521	.80	1043	.81
Supportive Relationships	3	1546	.81	1047	.79
Positive Social Norms	4	1536	.92	1060	.91
Support for Efficacy and Mattering	3	1525	.70	1023	.76
Opportunities for Skill Building	6	1510	.91	1015	.90
Program Quality (Total)	**22**	1390	.94	943	.94

Core Competencies Scale Reliability Estimates

Table 2. All scales had acceptable reliability estimates at pre- and post-surveys.

Core competencies are behaviors and attitudes related to five constructs that are common to the goals of all CYFAR SCPs: social conscience, personal values, caring, decision-making, and critical thinking. This measure contains 30 questions. There is only one version of this measure, directed toward youth.

All measures of core competencies demonstrated high reliability at pre-survey and post-survey (see Table 2). The measure of social conscience demonstrated the highest reliability, while the measure of caring demonstrated the lowest reliability, though all reliabilities were adequate.

Table 2. Reliability Estimates for Core Competency Scales					
Scale	**Items**	**Pre-Survey**		**Post-Survey**	
		n	**α**	***n***	**α**
Social Conscience	6	2624	.84	1374	.90
Personal Values	5	2481	.80	1371	.88
Caring	9	2390	.80	1306	.80
Decision Making	5	2531	.83	1355	.85
Critical Thinking	5	2561	.81	1359	.86

Common Measures of Short-Term Outcomes Reliability Estimates

There are 13 instruments related to short-term outcomes from which SCPs can choose the measure or measures that best fit their program goals. These measures fit into seven categories: nutrition, physical activity, leadership development, technology, science, workforce preparation, and parenting.

Leadership development. Leadership development is a 9-item instrument measuring self-efficacy for civic engagement in high-school youth.

Nutrition. There are two versions of the nutrition measure, a 17-item instrument aimed at youth in high school, and a 9-item instrument aimed at youth in grade three through middle school. These instruments measure youth's self-efficacy regarding sustaining healthy eating behaviors and eating healthy foods.

Parenting. This 18-item instrument measures attitudes and behaviors related to child-rearing. There is one version of this measure, directed at adults.

Science. Science is measured through two instruments: a 23-item instrument that measures attitudes toward math and science and science self-efficacy for youth in middle and high school, and a 13-item instrument that measures attitudes toward science for youth in grades three through five.

Technology. There are three versions of the technology common measure: an 8-item instrument measuring self-efficacy toward computers for adults, a 10-item instrument measuring ease of use and skill-level with computers for youth in middle and high school, and a 9-item measure of computer importance and enjoyment for youth in grades three through five.

Workforce preparation. There are two versions of this measure of effective problem-solving and decision-making behaviors and perceptions of competence: a 15-item version for high-school youth and a 9-item version for middle-school youth.

Physical activity. There is one version of this measure, aimed at youth in middle and high school. It is a 3-item instrument that measures physical activity and screen time.

SCP evaluators use Survey Builder to choose from the above measures to create the survey that is most applicable to their SCP. After building the survey, they are able to use a paper-and-pencil or online version of the survey with SCP participants. Data are then imported into Survey Builder either through the online survey or through an upload of a .csv (comma separated values file). These data can then be exported or used to build a report that summarizes the impact of the SCP's work.

Table 3. Reliability estimates for common measures of short-term outcomes ranged from low to high depending on the measure. The technology and science scales for grades three through five generally obtained lower reliability estimates than similar scales for middle and high school youth.

Table 3. Reliability Estimates for Common Measures of Short-Term Outcomes

Scale	Items	Pre-Survey		Post-Survey	
		n	α	*n*	α
Leadership Development	9	751	.91	512	.93
Nutrition (High School)	9	694	.84	389	.83
Nutrition (Middle School)	9	217	.80	149	.83
Nutrition (Grades 3 to 5)	9	441	.79	364	.77
Parenting	18	498	.91	178	.89
Science (MS/HS)	23	112	.93	35	.89
Technology (Adults)	8	58	.89	39	.86
Technology (MS/HS)	10	57	.90	1	-

Workforce Prep (HS)	15	227	.83	160	.85
Workforce Prep (MS)	9	859	.62	483	.69
Physical Activity (Grades 3 to 5)	11	114	.83	71	.88
Science (Grades 3 to 5)	13	431	.81	311	.74
Technology (Grades 3 to 5)	9	488	.63	438	.67
Note: Due to a low post-survey sample size, reliability for the Technology MS/HS scale was not calculated					

The sample sizes of specific common measures vary greatly among measures because SCPs choose only the measure or measures that are most applicable to their specific program. The reliability estimates for the common measures of short-term outcomes were variable, ranging from low to high (see Table 3). Notably, the technology- and science-related scales for middle school and high school youth had relatively good reliability estimates, but the science and technology scales for grades three through five generally obtained lower reliability estimates. Further evaluation of the science and technology scales used for grades three through five will be considered to determine what factors might be affecting the reliability estimates.

Other measures demonstrated high reliability, including: leadership development, parenting, middle- and high-school science, and middle- and high-school technology. The middle-school workforce preparation measure demonstrated less reliability than the high-school workforce preparation measure. It may be useful to examine factors contributing to the lower reliability of the middle-school workforce preparation measure. Reliability estimates were similar across pre- and post-surveys for all measures.

Data Collection

Obtaining data from every program is key to evaluating the outcomes of the CYFAR Initiative. CYFAR data collection began in 2011 and has continued for the past five years. Though there are certain requirements guiding what types of measures SCPs need to use in their questionnaires, actual use has varied over time. In the time from 2011 through 2016, however, the number of SCPs collecting data has increased substantially, with a significant spike in 2015 after data collection became mandatory in 2014. In 2016, the number of SCPs collecting data has continued to increase, reaching a high of 25 SCPs submitting data in the past year.

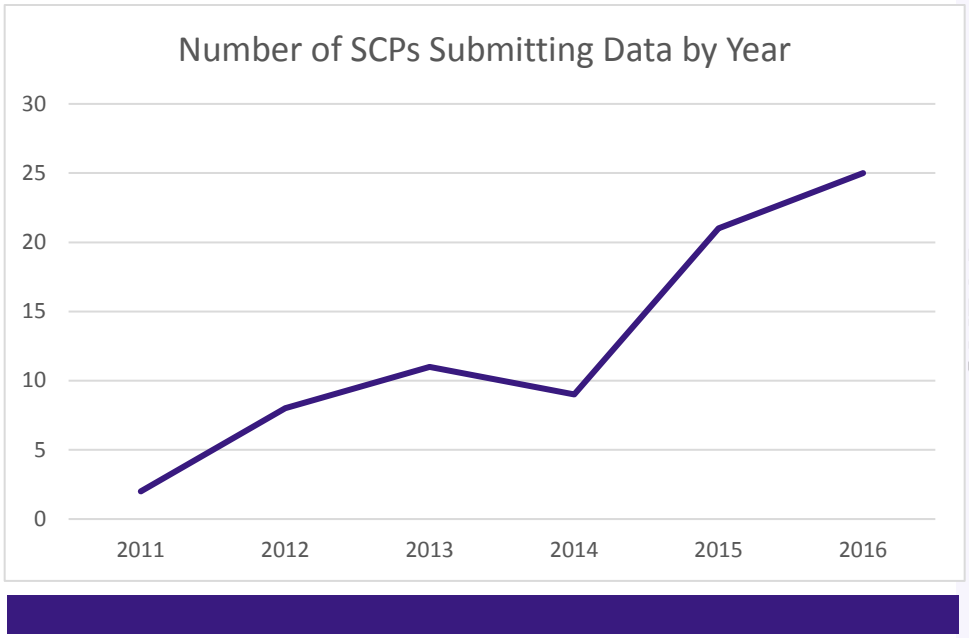

Figure 1. The number of SCPs submitting data through the CYFAR Survey Builder tool has increased over time, significantly so in the past two years.

The number of participants from whom data are collected has shown a similar pattern, rising from 23 participants in 2011 to a high of 3,350 participants in 2016. Here, too, collection of data significantly increased in 2015 and has continued to increase since.

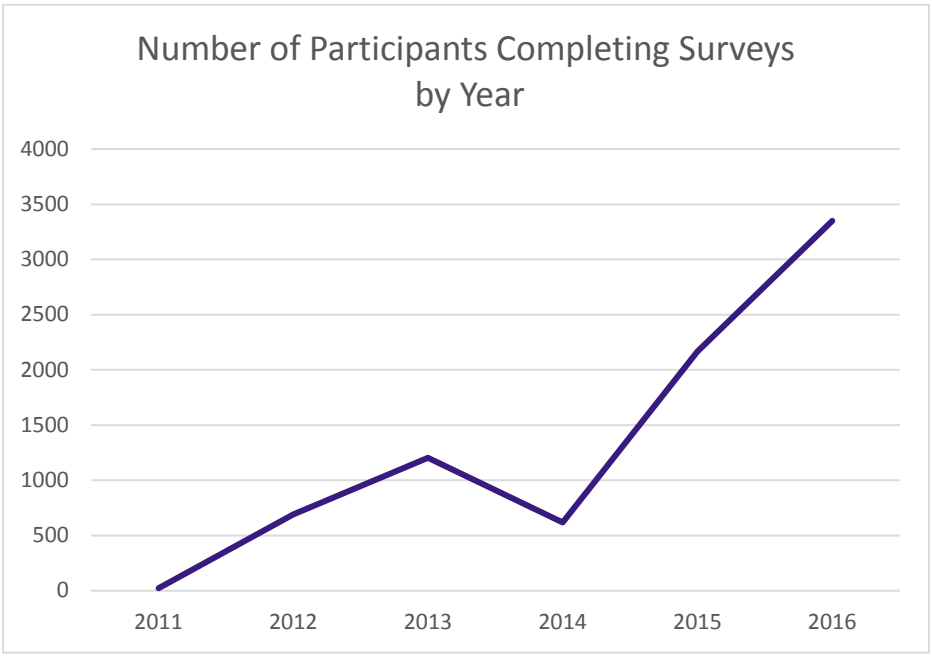

Figure 2. The number of participants for whom data has been submitted through the CYFAR survey builder tool has also increased over time, particularly over the past two years.

Within the pre- and post-survey design of CYFAR data collection, the collection of pre- and post-survey data are increasing over time, though, every year, there continue to be more pre-surveys submitted than post-surveys. Note that not all post-surveys submitted have a matching pre-survey. Youth participating in any given SCP at pre-survey may not be the same at post-survey. Furthermore, a significant proportion of SCPs are not using the requested naming convention to develop participants' unique identifiers (more than 24% of participants have an identifier that clearly does not follow the naming convention), and it is not possible to pair pre- and post-surveys when SCPs do not follow the naming convention.

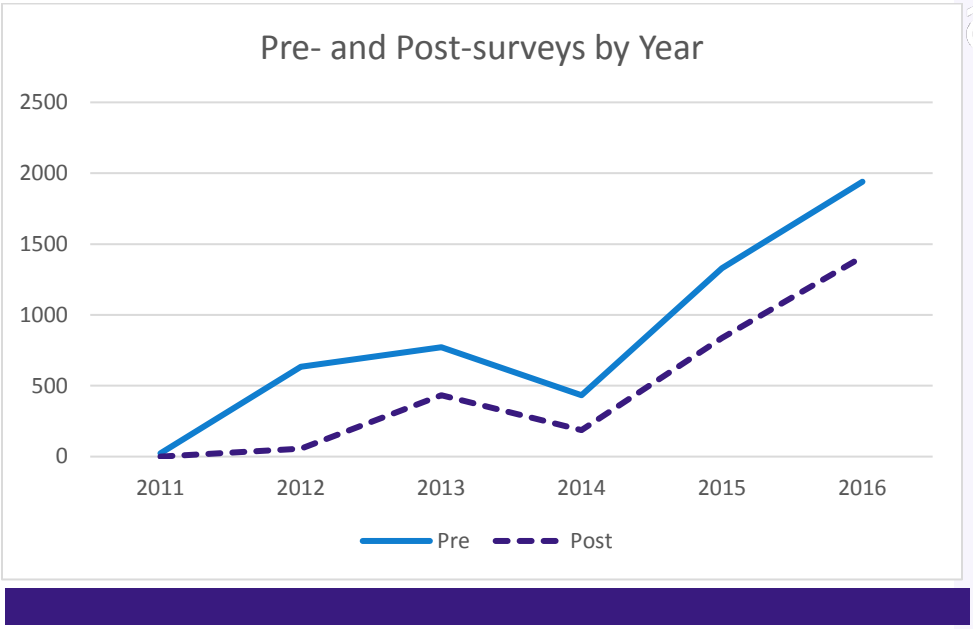

Figure 3. Collection of pre-surveys and post-surveys are both increasing over time; however, there continue to be more pre-surveys collected each year than post-surveys.

There does seem to be some confusion regarding the use of pre- and post-surveys. There have been several situations in which the title of a survey conveys that it is a post-survey and all of the responses are coded as a pre-survey. It is possible, given that the post-survey does not allow for collection of demographic data, that SCPs are using the pre-survey as a post-survey; however, there is no way to indicate this in the CYFAR data collection system.

Types of Data Collected

Overall collection of any type of data is increasing. Examining patterns of collection of each different type of data can give more nuanced insight into this pattern, demonstrating which measures are most likely to be used and how that is changing over time.

Program quality. As noted previously, collection of program quality data are required for all SCPs with middle- and high-school youth. Collection of these data is optional for SCPs with elementary school youth and adults.

Collection of program quality data for middle- and high-school youth has risen steadily since 2011. Furthermore, the number of participants who have not completed the measure of program quality (Program Quality Instrument; PQI) in this age group has remained somewhat low with the exception of data collected in 2012. The proportion of middle- and high-school youth completing the measure has been between 70% and 80% for the last three years.

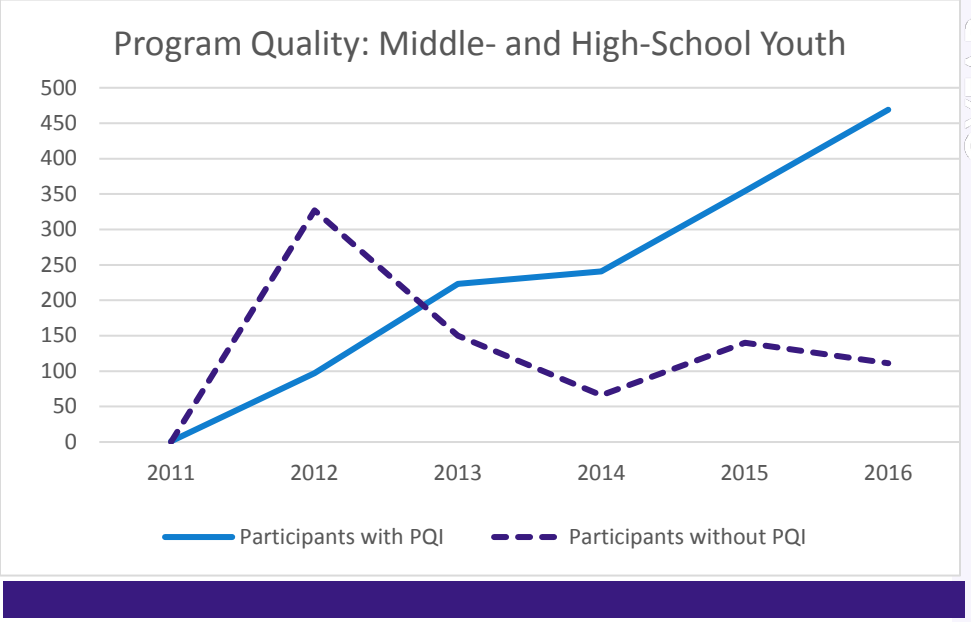

Figure 4. The collection of program quality data has risen over time, and, for the most part, the proportion of middle- and high-school youth completing the measure has remained high.

For youth in third through fifth grades, however, completion of the PQI is optional, and thus collection of program quality data in that group follows a different pattern. The increase in youth completing the PQI in this age group has been steady over time, while the number of youth not completing the PQI in this age group has been less uniform, with a dip in 2014 and significant increases in 2012 and 2015. In most years, the proportion of youth in grades three through five completing a measure of program quality has been under 50%.

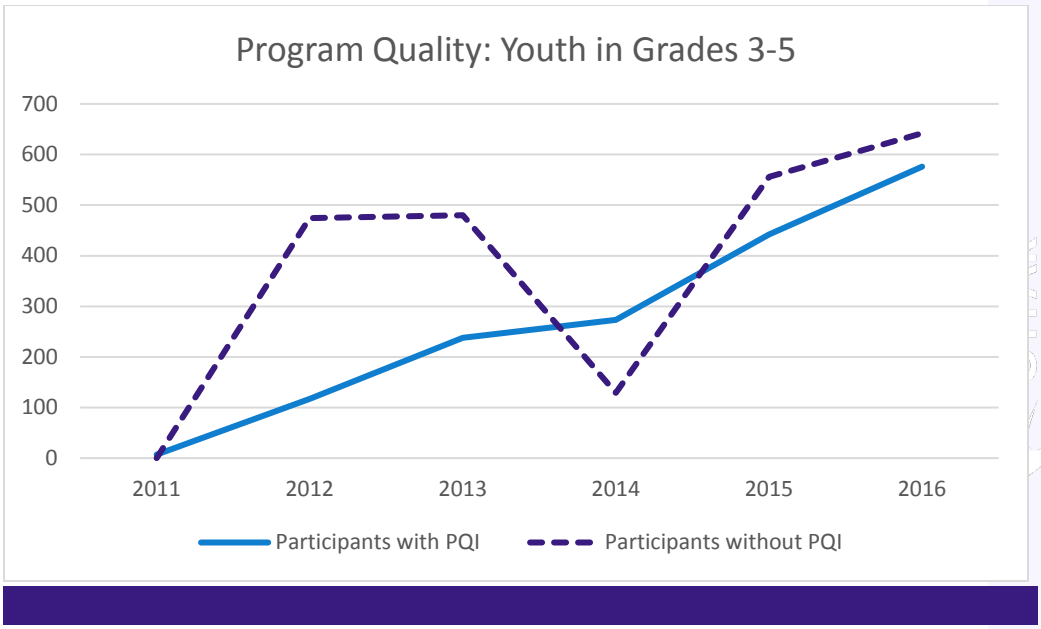

Figure 5. The number of youth in grades three through five completing the measure of program quality has increased steadily over time. In most years, less than half of these youth have completed this measure.

The measurement of program quality is increasing over time, particularly for youth in middle and high school. However, the proportion of youth in grades three through five not completing the PQI is also increasing. Given that these youth make up a significant proportion of the CYFAR participants completing data collection (over a third), it could be useful to investigate why this is the case.

Core competencies. The measures of core competencies follow the same requirements as the measure of program quality; that is, they are required for middle- and high-school youth, but optional for youth in grades three through five. The collection of these data also follows a similar pattern to the collection of program quality data.

Almost all youth in middle and high school in the CYFAR data set have completed the measures of core competencies. The number of middle- and high-school youth who have completed these measures has increased over time, while the number of middle- and high-school youth who have not completed these measures has remained fairly constant, at less than 5%.

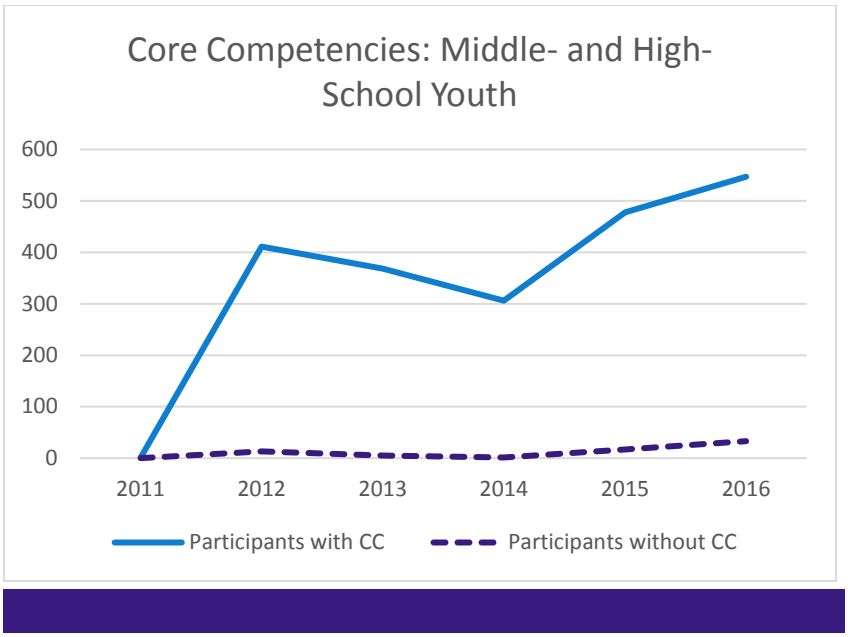

Figure 6. The number of middle- and high-school youth completing measures of core competencies has increased over time, while the number not completing these measures has remained fairly constant and very low.

For youth in grades three through five, however, both the number of youth completing the measures and the number of youth not completing the measures has increased over time. Each year, the proportion of youth in grades three through five completing measures of core competencies has been less than 45%.

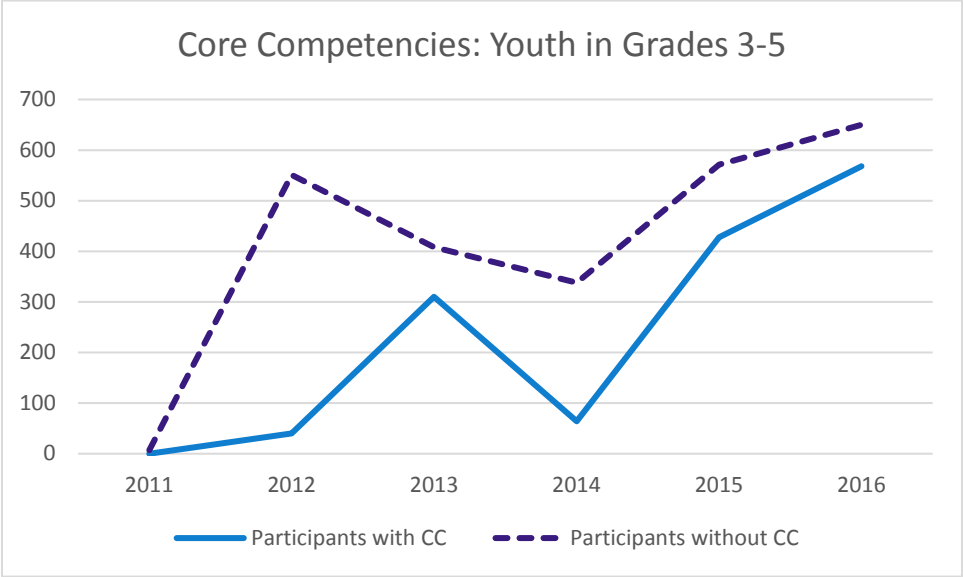

Figure 7. The number of youth in grades three through five completing measures of core competencies has increased over time; however, less than half of these youth have completed the measures at any given time.

As with program quality measurement, measurement of core competencies has increased significantly over time. Among middle- and high-school youth, most complete the measures of core competencies; however, among youth in grades three through five, the proportion completing these measures has remained low. Again, given that this group is a significant portion of the overall number of CYFAR participants completing data collection, it could be useful to investigate this phenomenon further.

Common measures. As noted previously, SCPs are able to choose from seven common measure areas, which contain measures of short-term outcomes that are relevant to each SCP. Many SCPs choose to use more than one common measure (e.g., measuring nutrition and physical activity or leadership development and nutrition). Among data collected since 2011, nutrition is the most frequently chosen common measure to include in CYFAR surveys, with 34% of participants completing this measure. This is followed closely by workforce preparation, with 23% of all participants completing that measure.

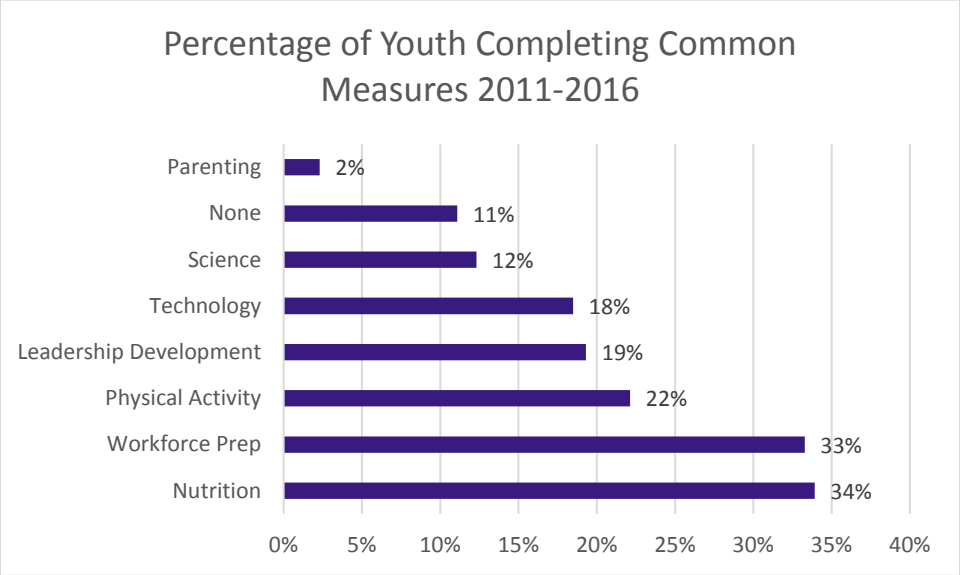

Figure 8. Workforce preparation is the most-used common measure of short term outcomes, followed closely by nutrition.

The use of all common measures of short-term outcomes has increased over time. Furthermore, the combinations of common measures occurring most often are leadership development, nutrition, and physical activity, and leadership, science, and workforce preparation, with over 20% of participants completing one of these combinations of three measures.

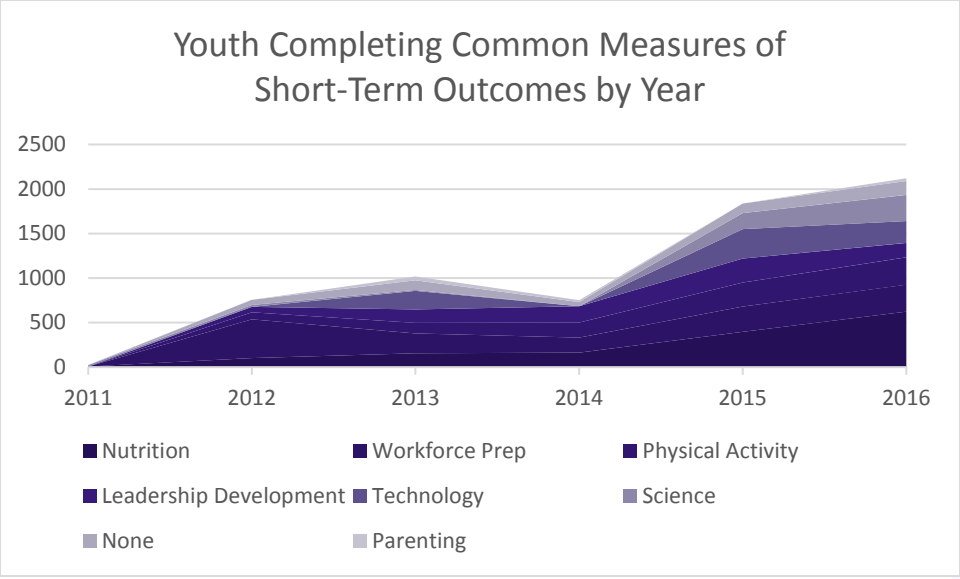

Figure 9. The number of youth completing common measures of short-term outcomes is increasing over time.

Most youth (74%) complete one or two common measures, though 16% complete more than two common measures. Notably, 11% of youth do not have data for any common measure. However, of those youth who do not have data for any common measure, 80% are in grade two or below, for whom there are no common measures available.

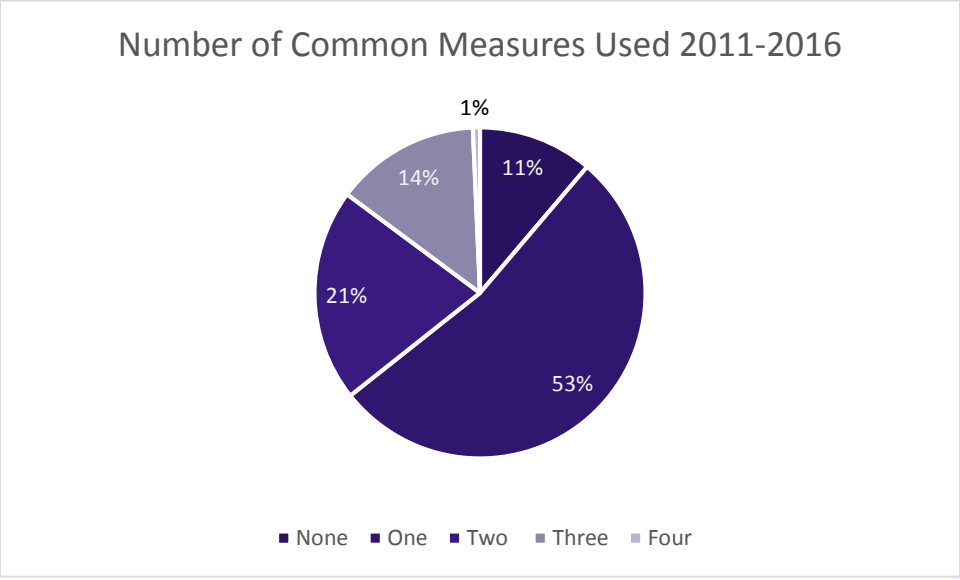

Figure 10. Most youth (74%) complete one or two common measures.

Parenting is a particularly interesting measure, with 58% of adult SCP participants completing it, in addition to the measure being used in a program for teen parents. Furthermore, the percentage of adults completing the parenting measure has been relatively high over time.

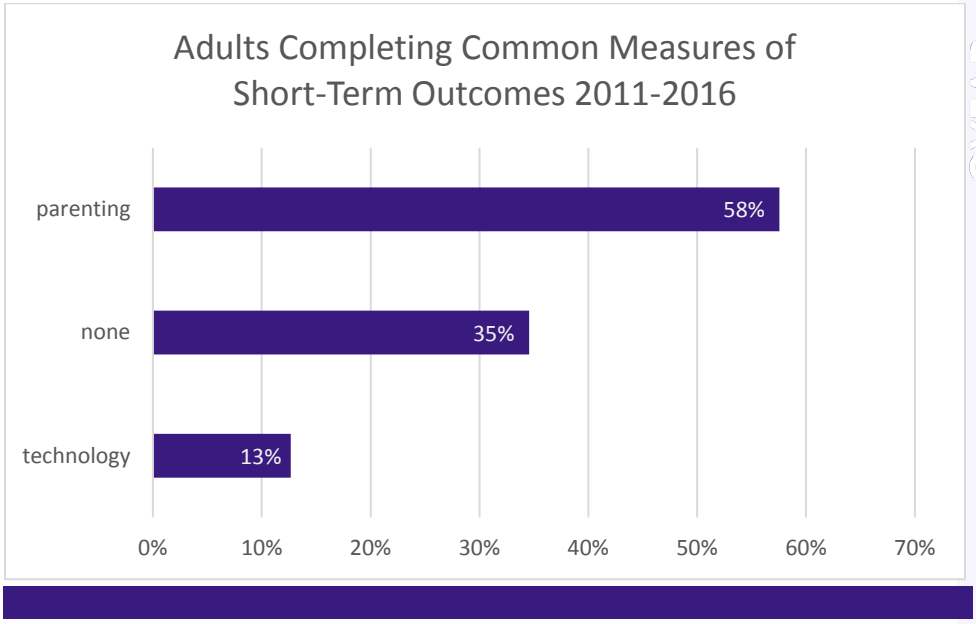

Figure 11. The majority of adults (58%) complete the parenting common measure.

Taken together, these data indicate that common measures are being used and that their use is increasing over time. Furthermore, workforce preparation is consistently the most-used common measure, followed closely by nutrition, with almost half of all youth completing one of those two measures.

In order to make recommendations regarding changes that can enhance CYFAR data collection efforts, it is useful to analyze the available data to gain a deeper understanding of the instruments, their relationships to one another, and what conclusions can be made regarding the potential impact of CYFAR SCPs.

Participants

This evaluation offers an understanding of those individuals who participate in the CYFAR initiative. Between 2011 and 2016, a total of 5,132 pre-surveys were completed: 4,259 youth pre-surveys and 813 adult pre-surveys. Demographic data has only been collected during pre-surveys.

Youth Population Served

Gender. Among the youth who completed CYFAR surveys, approximately 47% were males and 53% were females (see Figure 12).

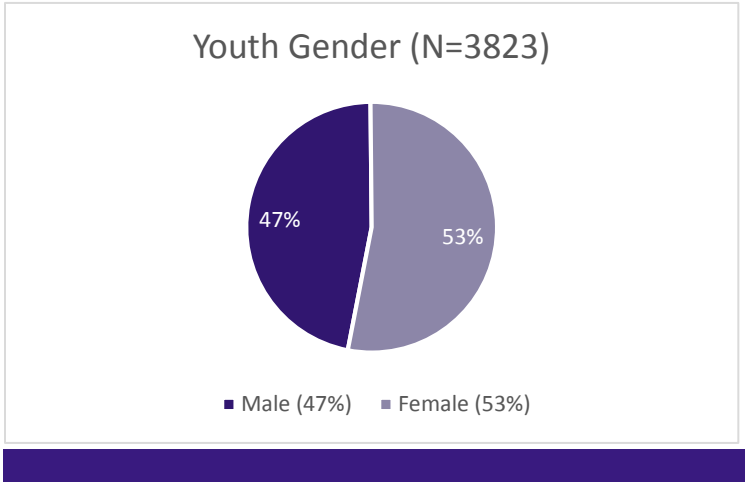

Figure 12. The percentage of surveys from female youth was slightly higher than male youth.

Table 4. Youth Grade Level		
Pre-Survey (*n*=3839)		
Grade Level	**Frequency**	**Percent (%)**
Pre-K	126	3%
K-2nd	78	2%
3rd-5th	1195	31%
6th-8th	1289	34%
9th-12th	1113	29%
Post-HS	38	1%

Table 4. The highest percentage of youth participants were in the 6th-8th grade category (33.6%).

Grade Level. The majority of youth in this sample were in grades three through twelve, with the highest proportion in grades six through eight (33.6%, see Table 4). A small number of youth respondents (1%) reported having graduated from high school, as some CYFAR programs serve youth into early adulthood.

Ethnicity and Race. Of CYFAR youth completing surveys, 36% of youth identified as Hispanic or Latino compared to 64% identifying as Non-Hispanic or Non-Latino. In addition, the largest percentage of youth identified as either White (52%) or Black (32%; see Figures 13 and 14)

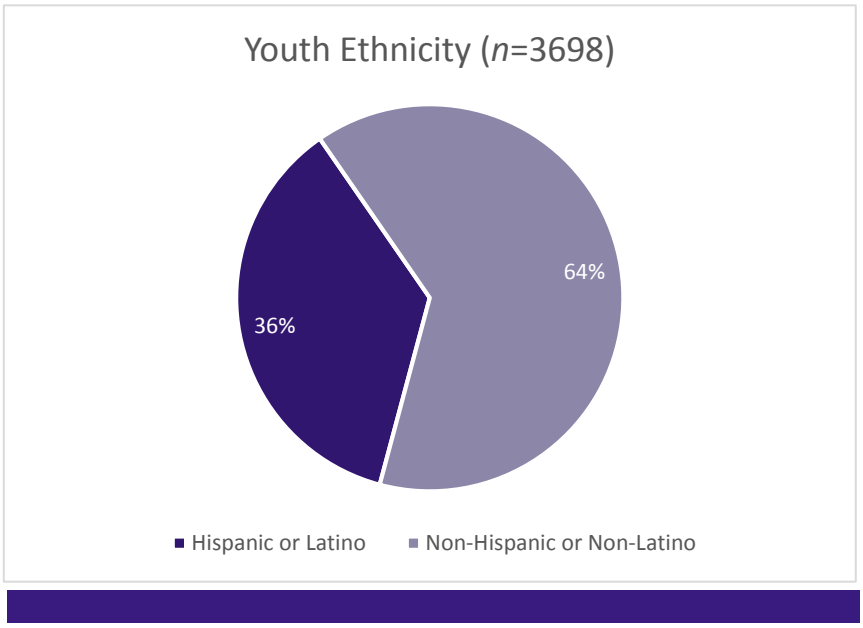

Figure 13. Youth identifying as Hispanic or Latino made up slightly more than one third of the youth surveys collected.

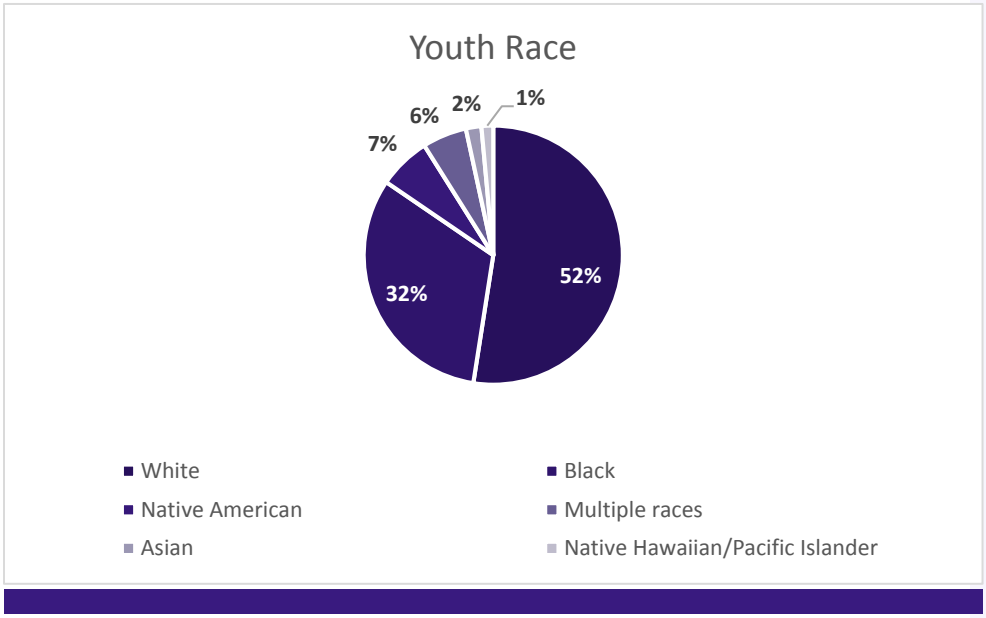

Figure 14. The largest percentage of youth respondents in CYFAR programs identified as White (52%) or Black (32%).

Military Parent. Of those youth participating in the survey, 8% indicated they had a parent in the military, most commonly in the Army (30%; see Table 5).

Table 5. Parents in the Military		
Pre-Survey (*n*=4115)		
Parents In the Military	**Frequency**	**Percent (%)**
Yes	327	8%
No	3788	92%
Branch (*n*=165)		
Air Force	47	14%
Army	95	30%
Guard	20	6%
Marine	31	10%
Navy	35	10%
Reserve	23	7%
Multiple Branches	46	14%
Not Specified	30	9%

Table 5. About 8% of youth reported having a parent in the military, most commonly in the Army (29%).

Adult Demographics

Adult Gender. There were more female (65%) than male (35%) adult respondents. The difference between the number of female and male participants was greater among adult participants than among youth participants.

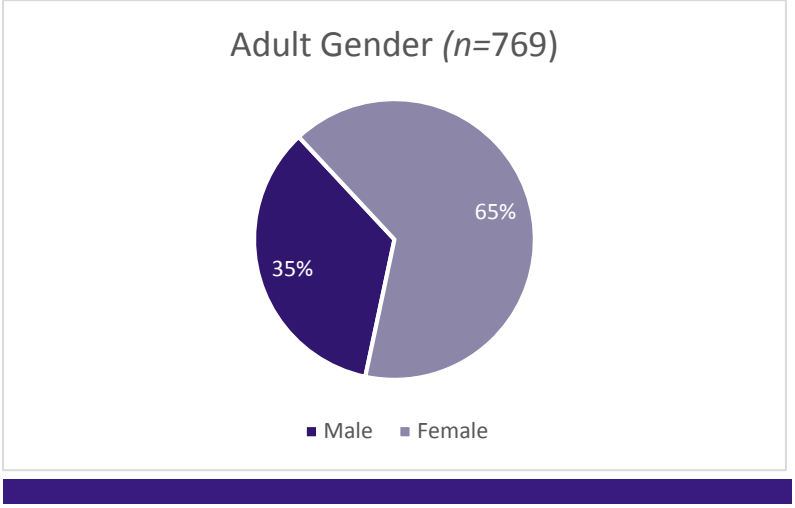

Figure 15. A much larger percentage of adult survey respondents were female than male.

Adult Ethnicity and Race. Of those adults completing surveys, 51% identified as Hispanic or Latino and 49% identified as Non-Hispanic or Non-Latino. A greater percentage of adult respondents identifed as Hispanic or Latino (51%) than in the youth demographic (36%). Similar to youth, the majority of adult participants also identified as either White (71%) or Black (22%).

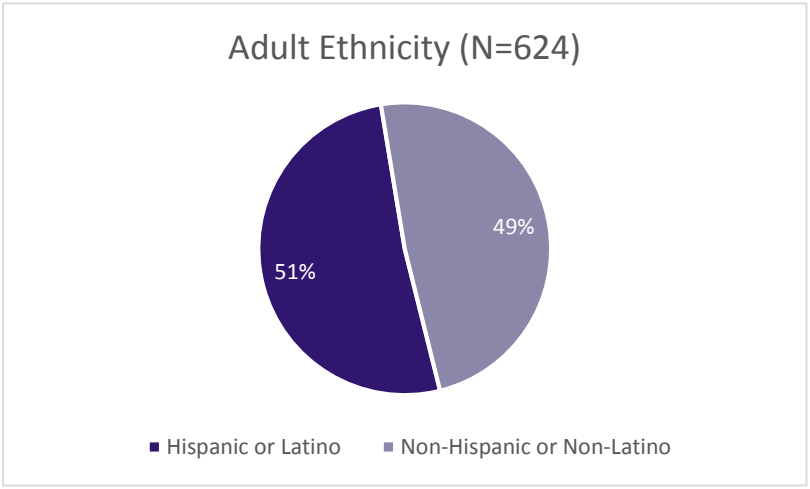

Figure 16. A similar percentage of adult survey respondents identified as Hispanic or Latino compared to those identifying as Non-Hispanic or Non-Latino.

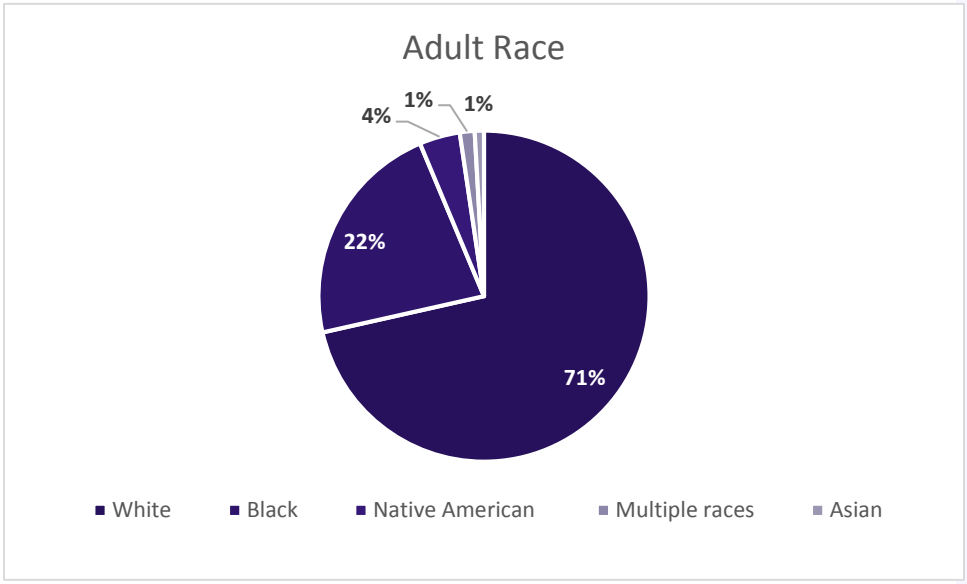

Figure 17. The majority of adult participants identified as White (71%) or Black (22%).

Core Competencies and Demographics

Examining demographic differences in how participants respond to CYFAR common measures is one way to gain insight into the individuals who participate in CYFAR SCPs. This exploration also allows for a deeper understanding of the common measures used and factors that may influence responses.

Table 6. Female youth scored significantly higher on all five Core Competency measures.

Gender. Female respondents scored significantly higher on all core competencies (social conscience, personal values, caring, decision-making, and critical thinking). Gender may influence how respondents experience CYFAR programming, or how individuals interpret items within the measures of core competencies. Another possibility is that there are differences between males and females in the general population on the constructs measured, and that these differences are evident within the population of CYFAR SCP participants as would be expected.

Table 6. Core Competencies Total Scale Score by Gender

Core Competency	Males		Females		Significance (*p*)
	n	Mean	*n*	Mean	
Social Conscience	1079	19.75	1273	21.10	<.001**
Personal Values	1014	16.69	1195	17.78	<.001**
Caring	967	27.52	1155	29.84	<.001**
Decision-Making	1073	14.73	1234	15.17	.002*
Critical Thinking	1079	13.79	1256	14.36	<.001**

* Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.

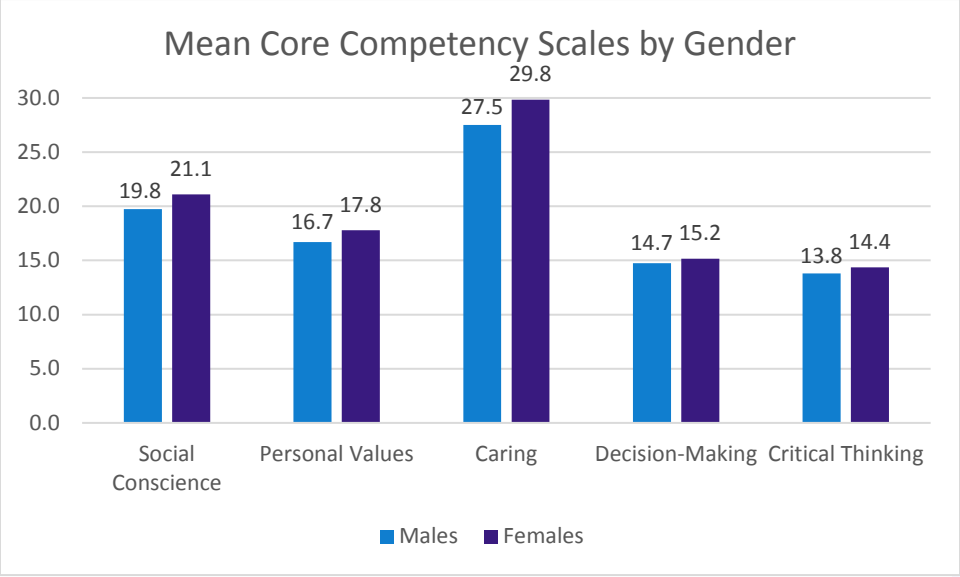

Figure 18. Females scored significantly higher than males on all Core Competency Scales.

Ethnicity. With regard to ethnic differences on measures of core competencies, Non-Hispanic or Non-Latino participants scored higher on most scales (with the exception of Personal Values) than Hispanic or Latino participants; however, this difference was statistically significant only for the Critical Thinking scale (see Table 7).

Table 7. Youth identifying as Non-Hispanic or Non-Latino scored significantly higher on the Critical Thinking Core Competency scale.

Table 7. Core Competencies Total Scale Score by Ethnicity					
Core Competency	**Hispanic or Latino**		**Non-Hispanic or Non-Latino**		**Significance (*p*)**
	n	**Mean**	***n***	**Mean**	
Social Conscience	920	20.36	1387	20.56	.17
Personal Values	788	17.38	1381	17.23	.251
Caring	756	28.50	1328	28.94	.075
Decision-Making	920	14.96	1344	15.00	.766
Critical Thinking	928	13.80	1363	14.35	<.001**
* Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.					

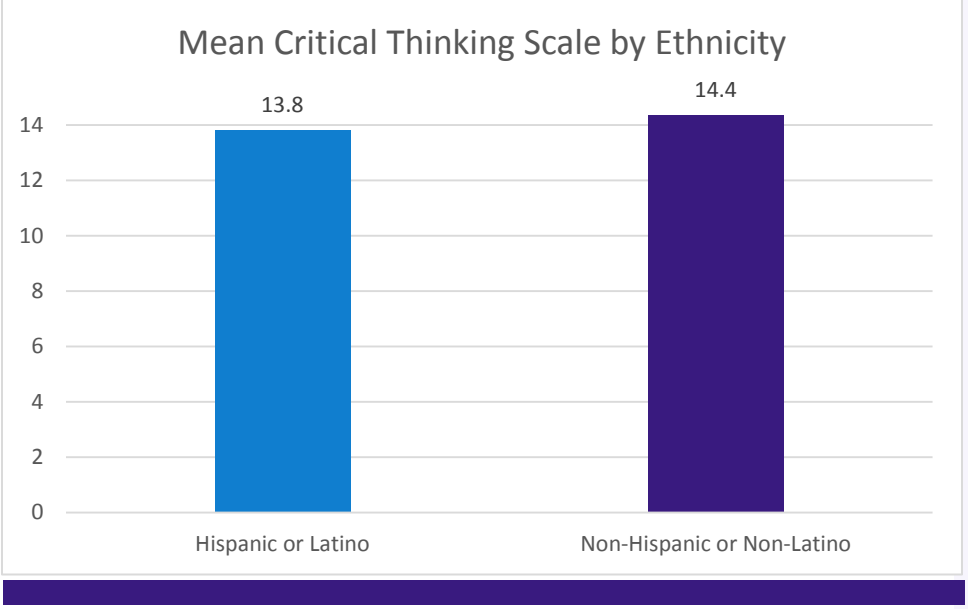

Figure 19. The Non-Hispanic or Non-Latino group scored significantly higher than the Hispanic or Latino group on measure of critical thinking.

Table 8. Youth identifying as Black scored significantly lower than youth identifying as White on the Caring scale, and significantly higher on the critical thinking scale.

Race. Statistically significant differences among races were found on measures of social conscience, caring, and critical thinking (see Table 8). Further analyses indicated that White youth scored higher on the measure of caring than Black youth. Furthermore, on the measure of critical thinking, Black youth scored higher than White youth. However, these results should be interpreted with caution due to large differences in sample sizes. Finally, despite evidence of statistically significant differences among races on the measure of social conscience, further analyses did not identify statistically significant differences between any two particular racial groups.

Table 8. Core Competencies Total Scale Score by Race

Core Competency	Mean						Significance (*p*)
	Native American	Asian	Black	Native Hawaiian or Pacific Islander	White	Multiple Races	
Social Conscience	19.84	19.56	20.77	21.70	20.51	20.77	.005*
Personal Values	16.79	16.76	17.33	17.94	17.24	17.50	.174
Caring	28.86	28.50	28.14[a]	28.42	29.1[a]	29.75	.010*
Decision-Making	14.94	15.63	15.24	15.27	14.92	14.99	.430
Critical Thinking	13.89	13.82	14.62[a]	14.50	14.01[a]	14.54	.009*

Note: [a] Denotes post-hoc test indicated means were significantly different at $p<.05$ level. * Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.

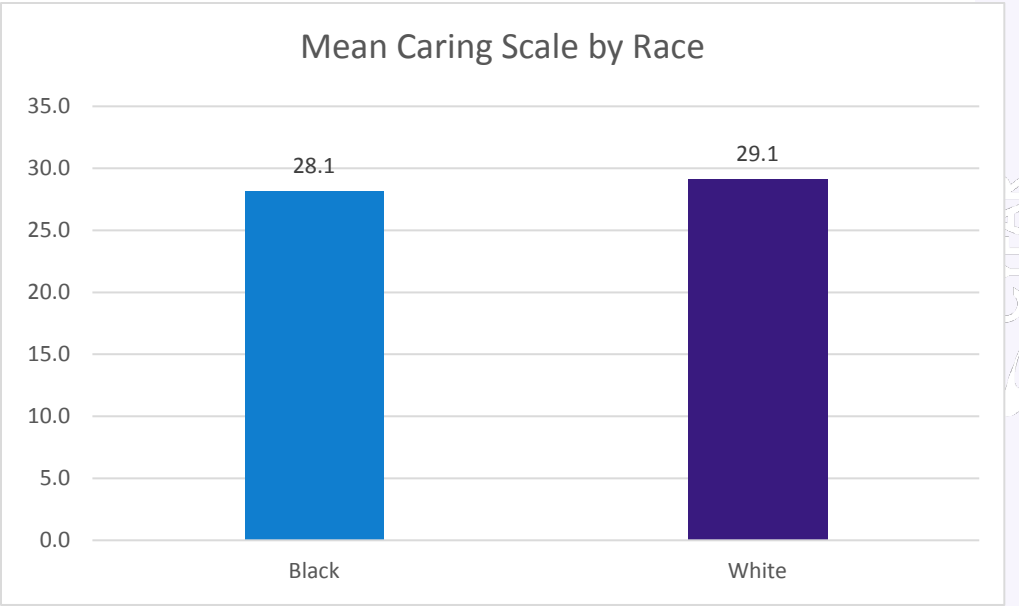

Figure 20. Youth identifying as White scored significantly higher on the measure of caring than youth identifying as Black.

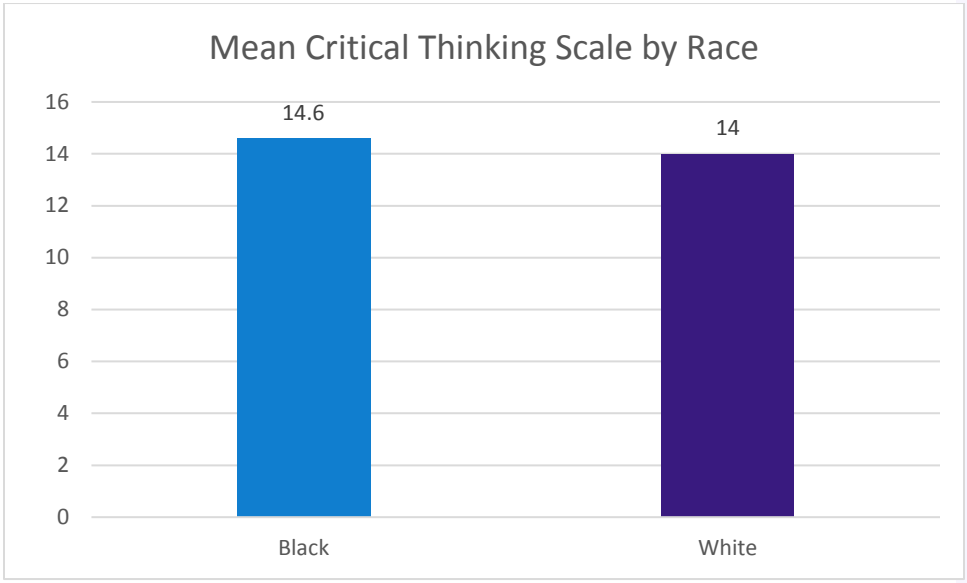

Figure 21. Youth identifying as Black scored significantly higher on the measure of critical thinking than youth identifying as White.

Common Measures of Short-Term Outcomes and Demographics

Common measures of short-term outcomes are related to the specific programmatic focus of SCPs. The 13 common measures of short-term outcomes address seven specific outcome areas: leadership development, nutrition, parenting, science, technology, workforce preparation, and physical activity.

Table 9. Female youth scored significantly higher than male youth on the leadership development and high-school nutrition measures. Male youth scored higher than female youth on the middle- and high-school science measure.

Gender. Analysis of gender differences in common measures of short-term outcomes revealed that females scored significantly higher than males on the leadership development and high-school nutrition measures, whereas males scored higher on the middle- and high-school science measure (see Table 9).

Table 9. Common Measure Total Scale Score by Gender

Common Measure	Males		Females		Sig. (*p*)
	n	Mean	*n*	Mean	
Leadership Development	264	30.05	394	32.78	*P*<.001**
Nutrition (High School)	237	22.31	349	23.57	*P*<.001**
Nutrition (Middle School)	104	20.49	102	21.23	0.223
Nutrition (Grades 3 to 5)	202	21.53	203	22.26	0.083
Parenting	16	90.69	65	90.74	0.992
Science (HS and MS)	36	70.06	48	61.31	0.003*
Technology (HS and MS)	19	31.05	24	32.00	0.670
Workforce Preparation (HS)	81	45.26	152	45.52	0.774
Workforce Preparation (MS)	425	26.14	362	26.06	0.937
Physical Activity (Grades 3 to 5)	34	27.68	41	25.46	0.059
Technology (Grades 3 to 5)	236	5.34	250	5.38	0.472
Science (Grades 3 to 5)	202	10.92	218	10.93	0.966

Note: Sample sizes (*n)* varied greatly between scales and between gender groups. * Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.

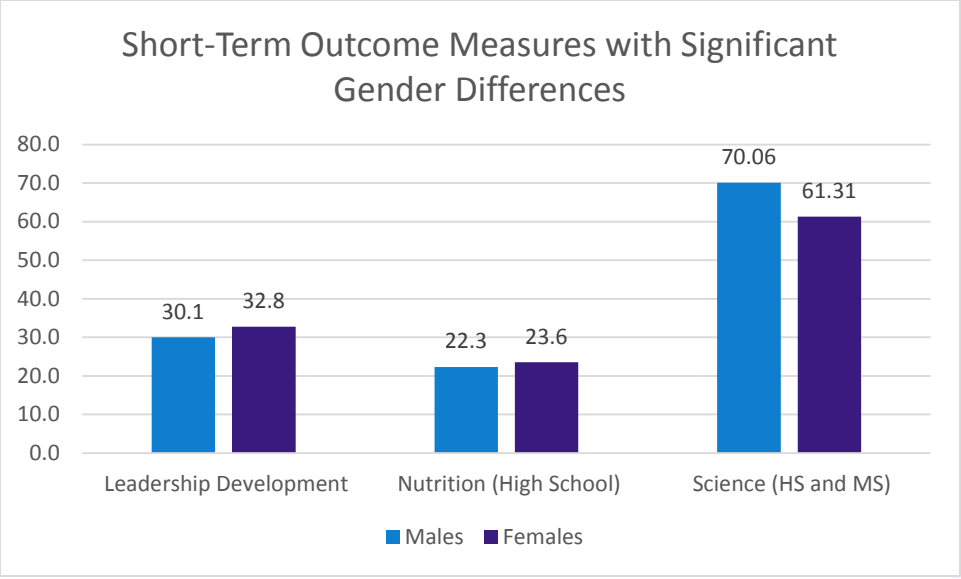

Figure 22. Female youth scored significantly higher than male youth on the leadership development and high-school nutrition measures. Male youth scored higher than female youth on the middle- and high-school science measure.

Table 10. In general, there were no differences on common measures of short-term outcomes between ethnicity groups with the exception of the measure of middle-school workforce preparation.

Ethnicity. In addition, there were differences between ethnicities on common measures of short-term outcomes. Middle-school youth identifying as Non-Hispanic or Non-Latino scored higher on the measure of workforce preparation than middle-school youth identifying as Hispanic or Latino (see Table 10). However, this difference was not present on the measure of high-school workforce preparation. It is also notable that the number of youth of different ethnicities varied greatly among the measures. This may be an indication of CYFAR SCPs tending to have different focuses depending on the population of youth they serve.

Table 10. Common Measure Total Scale Score by Ethnicity

	Hispanic or Latino		Non-Hispanic or Non-Latino		
Common Measure	***n***	**Mean**	***n***	**Mean**	**Sig. (p)**
Leadership Development	386	31.33	263	32.08	0.233
Nutrition (High School)	431	23.19	151	22.76	0.224
Nutrition (Middle School)	54	21.07	151	20.79	0.679
Nutrition (Grades 3 to 5)	209	21.75	194	22.12	0.376
Parenting	22	88.64	57	91.33	0.560

Science (HS and MS)	26	66.42	56	64.34	0.520
Technology (HS and MS)	12	33.58	30	30.70	0.245
Workforce Preparation (HS)	44	45.34	183	45.34	0.998
Workforce Preparation (MS)	264	25.40	508	26.36	0.002*
Physical Activity (Grades 3 to 5)	16	27.50	57	26.37	0.427
Technology (Grades 3 to 5)	251	5.31	230	5.41	0.101
Science (Grades 3 to 5)	23	10.49	381	10.96	0.341
* Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.					

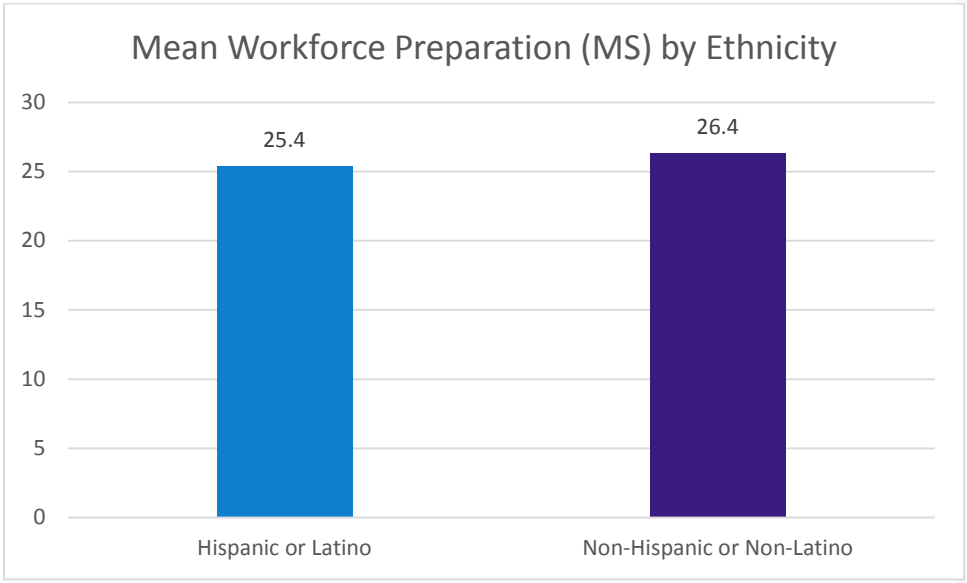

Figure 23. Non-Hispanic or Non-Latino youth scored significantly higher on the measure of middle-school workforce preparation than Hispanic or Latino youth.

Table 11. In general, there were no significant differences between racial groups on the common measures of short-term outcomes.

Race. In general, there were few significant differences based on race across the common measures of short-term outcomes (see Table 11). Overall analyses indicate a significant difference among races on the leadership development scale; however, further analyses did not indicate a significant difference between any two racial groups, which may be due to large variability in the number of youth of each race completing the measure.

Table 11. Common Measures of Short-Term Outcomes by Race

Common Measure	Mean						Sig. (*p*)
	Native American	Asian	Black	Native Hawaiian or Pacific Islander	White	Multiple Races	
Leadership Development	27.81	29.56	33.33	29.43	31.73	27.81	.008*
Nutrition (High School)	23.76	20.79	22.12	24.70	23.27	23.76	.069
Nutrition (Middle School)	21.60	21.60	20.30	14.33	21.47	21.60	.072
Nutrition (Grades 3 to 5)	20.97	22.50	20.73	27.00	21.78	20.97	.434
Parenting	91.00	-	90.95	-	92.48	91.00	.256
Science (HS and MS)	57.80	78.50	65.63	55.00	68.73	57.80	.092
Technology (HS and MS)	28.50	-	34.82	30.00	29.67	28.5	.237
Workforce Preparation (HS)	45.88	47.25	44.67	43.00	46.02	48.50	.185
Workforce Preparation (MS)	26.00	24.77	26.08	26.00	26.28	26.37	.833
Physical Activity (Grades 3 to 5)	27.75	-	26.67	-	26.49	25.38	.886
Technology (Grades 3 to 5)	5.23	5.31	5.33	5.38	5.42	5.28	.784
Science (Grades 3 to 5)	10.54	11.85	10.66	12.33	11.11	10.81	.430

* Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.

Correlations between Core Competencies and Common Measures

Correlations determine whether there is a statistically significant relationship between two variables. They also convey whether the relationship is positive or negative (i.e., inverse) and the strength of the relationship. A statistically significant positive correlation indicates that, overall, participants who score higher on one measure also score higher on the other measure (though there is always variation at the individual level).

Pre-survey data were analyzed to determine the relationship between core competencies and common measures of short-term outcomes, which revealed several statistically significant correlations. Notably, there was a significant positive relationship between four common measures of short-term outcomes and all core competencies (see Table 12). These four common measures of short-term outcomes were: leadership development, high-school nutrition, high-school workforce preparation, and middle-school workforce preparation. Parenting was also positively related to all core competencies except for caring.

This indicates that at the beginning of participation in a CYFAR SCP, youth who score higher on the core competencies also tend to score higher on the common measures of short-term outcomes.

Interestingly, there were significant relationships between high-school nutrition and all of the core competencies, but virtually no significant relationships were found between middle-school nutrition and core competencies.

Table 12. There were significant correlations between the core competencies and several common measures of short-term outcomes.

Table 12. Correlations Between Core Competencies and Common Measures at Pre-Survey

Common Measure	Correlation				
	Social Conscience	**Personal Values**	**Caring**	**Decision Making**	**Critical Thinking**
Leadership Development	.382**	.427**	.364**	.415**	.528**
Nutrition (HS)	.138**	.195**	.146**	.276**	.201**
Nutrition (MS)	.006	.078	.092	.048	.018
Nutrition (Grades 3 to 5)	-	-	-	-	-
Parenting	.335**	.352**	.177	.425**	.394**
Science (HS and MS)	.145	.155	.056	.194*	.320**
Technology (HS and MS)	.259	.303*	.124	.468**	.656**
Workforce Preparation (HS)	.182**	.299**	.279**	.317**	.335**
Workforce Preparation (MS)	.279**	.278**	.310**	.326**	.306**
Physical Activity (Grades 3 to 5)	-	-	-	-	-
Science (Grades 3 to 5)	.185	.132	.314*	.346**	.452**
Technology (Grades 3 to 5)	.170	-.072	.184	.253*	.267*

Note: Low sample sizes prevented analysis of the Nutrition and Physical Activity measures for the 3rd-5th grade levels. * Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.

Program Quality

The measurement of program quality is fundamental to the mission of improving programming for individuals served in CYFAR SCPs. Furthermore, understanding the relationship between program quality and outcome measures helps determine how various aspects of program quality might be associated with positive youth development. Post-survey correlations between program quality and core competencies and between program quality and common measures of short-term outcomes help to illuminate the relationship between program quality and outcomes for CYFAR SCP participants. It is important to note that correlations do not imply cause and effect; they are only able to demonstrate a relationship between two variables, and they cannot specify a causal directionality to that relationship.

Table 13. All Program Quality scales correlated significantly with all Core Competency scales at post-survey.

Predictors of Core Competencies

All aspects of program quality were significantly correlated with all core competencies, such that higher scores on each aspect of program quality was associated with greater core competencies (see Table 13). Overall program quality was also significantly associated with all core competencies, such that higher overall program quality was associated with higher core competencies (i.e., social conscience, personal values, caring, decision-making, and critical thinking).

Table 13. Correlations Between Program Quality and Core Competencies at Post-Survey

	Program Quality Scale					
Core Competency Scale	**Safety**	**Relationships**	**Positive Social Norms**	**Support for Efficacy**	**Skill Building**	**Total Program Quality**
Social Conscience	.235**	.455**	.377**	.382**	.464**	.461**
Personal Values	.274**	.441**	.411**	.411**	.476**	.482**
Caring	.302**	.313**	.332**	.293**	.349**	.376**
Decision-Making	.185**	.355**	.355**	.362**	.398**	.400**
Critical Thinking	.182**	.361**	.346**	.380**	.387**	.405**

* Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.

Predictors of Common Measures of Short-Term Outcomes

The relationships between the program quality and common measures of short-term outcomes varied greatly among different scales (see Table 14). Analyses revealed a significant, positive relationship between program quality (all individual aspects and overall quality) and leadership development and workforce preparation. Youth who reported higher program quality tended to report higher leadership development and workforce preparation.

Table 14. There are relationships between youth perceptions of program quality and some of the common measures of short-term outcomes, such that youth who rate program quality higher experience more positive short-term outcomes (for a subset of outcomes measured).

The relationship between the program quality and nutrition was inconsistent, which makes firm conclusions difficult. For high-school youth, nutrition was associated with all aspects of program quality except for safety; however, for middle-school youth, nutrition was only associated with safety. Furthermore, for youth in grades three through five, nutrition was not significantly related to any aspect of program quality.

Generally, there is a relationship between youth perceptions of program quality and some of the common measures of short-term outcomes, such that youth who rate program quality higher experience more positive short-term outcomes (for a subset of outcomes measured). Low post-survey sample sizes may have hindered the ability to determine whether there are relationships between program quality and other short-term outcomes.

Table 14. Correlations Between Program Quality and Common Measures at Post-Survey						
	Program Quality Scale					
Common Measure	Safety	Relationships	Positive Social Norms	Support for Efficacy	Skill Building	Total Program Quality
Leadership Development	.308**	.400**	.439**	.510**	.472**	.500**
Nutrition (High School)	.074	.164**	.207**	.131**	.213**	.183**
Nutrition (Middle School)	.366**	.025	.062	.055	.118	.123
Nutrition (Grades 3 to 5)	-.057	.165	.056	.212	.061	.101
Parenting	-.186	-.221	-.038	-.258	-.222	-.216
Science (HS and MS)	.309	.218	.215	.049	.337	.288
Technology (HS and MS)	-	-	-	-	-	-
Workforce Preparation (HS)	.306**	.209**	.223**	.310**	.365**	.333**
Workforce Preparation (MS)	.256**	.343**	.191*	.397**	.403**	.407**
Physical Activity (Grades 3 to 5)	0.082	0.13	0.361	0.112	0.328	0.312
Science (Grades 3 to 5)	-	-	-	-	-	-
Technology (Grades 3 to 5)	-	-	-	-	-	-
Note: Low sample sizes for *Technology (HS and MS), Science (Grades 3 to 5),* and *Technology (Grades 3 to 5)* made analysis impossible. * Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.						

Current data are able to provide some insight into participants taking part in CYFAR SCPs. Generally, higher program quality is associated with more positive youth outcomes; however, as noted before, this does not imply causation, and the relationship between the two variables may be bi-directional. It is possible that youth who score higher on core competencies and short-term outcomes are more likely to rate program quality as higher, or that youth with higher levels of core competencies increase program quality through their contributions. It is also possible that higher quality programs cause more positive outcomes for youth.

Comparing Pre- and Post-Survey Outcomes

Due to inconsistencies in the way SCPs follow existing participant-naming and data-collection conventions, the pre-survey and post-survey scores were treated as independent groups in the following analyses. Because some of the data used for the analysis do include matched-pair scores, the true independence of the groups is compromised. There is also a large difference in sample size between pre-survey and post-survey (with a smaller number of participants completing the post-survey), which may affect comparisons between pre- and post-surveys. However, useful information can still be gained through analysis of differences in outcomes between pre- and post-surveys groups.

Table 15. Core Competency scale scores were significantly lower at post-survey for the Social Conscience and Personal Values scales, and significantly higher at post-survey for the Decision-Making and Critical Thinking scales.

Core Competencies

Some significant differences exist between the pre- and post-survey groups with regard to core competencies (see Table 15). The differences between the pre- and post-survey means are relatively small, so significant findings may have resulted from having a large sample size. Post-survey means were significantly lower than pre-survey means on the measures of social conscience and personal values. However, post-survey means were significantly higher than pre-survey means on the measures of critical thinking and decision-making. It is important to note that this analysis is examining pre-surveys and post-surveys as whole groups, rather than pairing data and looking at individual change from pre- to post-survey.

Table 15. Comparison of Core Competency Scale Score Pre-Survey to Post-Survey

Core Competency	Pre-Survey		Post-Survey		Sig. (*p*)
	n	Mean	*n*	Mean	
Social Conscience	2624	20.49	1374	20.15	.004*
Personal Values	2481	17.23	1371	17.02	.032*
Caring	2390	28.70	1306	28.91	.261
Decision-Making	2531	14.95	1355	15.20	.028*
Critical Thinking	2561	14.06	1359	14.36	.008*

* Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.

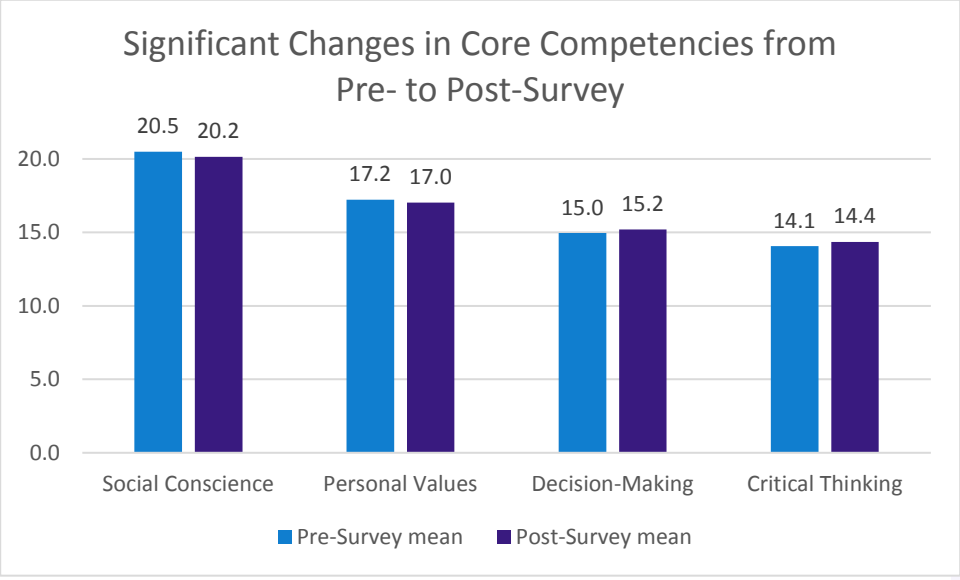

Figure 24. Measures of social conscience and personal values significantly decreased from the pre-survey to post-survey. Measures of decision-making and critical thinking significantly increased from pre-survey to post-survey.

Common Measures of Short-Term Outcomes

Sample sizes are much smaller for common measures of short-term outcomes because each SCP is only required to collect data for only one common measure of short-term outcomes, with most measures targeting specific age groups and program types. There were some significant differences between youth at pre-survey and youth at post-survey (See Table 16). Youth scored higher on six common measures of short-term outcomes at post-survey when compared to pre-survey: leadership development, high-school nutrition, parenting, high-school workforce preparation, technology for grades three through five, and science for grades three through five. The highest increases from pre-survey to post-survey were in leadership development and parenting. In addition, although scores were significantly higher on high-school nutrition at post-survey, the same pattern was not present for other age groups. Further investigation into factors that may influence the changes in scores of youth in different age groups may be useful for understanding the various programming needs of youth participating in CYFAR SCPs.

Table 16. Six Common Measure scales increased for the post-survey group compared to the pre-survey: Leadership Development; Nutrition (High School); Parenting; Workforce Preparation (High School); Technology (Third-Fifth grade); Science (Third-Fifth).

Table 16. Comparison of Common Measure Total Scale Score Pre-survey to Post-Survey

Common Measure	Pre-Survey		*Post-Survey*		Sig. (*p*)
	n	**Mean**	***n***	**Mean**	
Leadership Development	751	31.31	512	33.43	*p*<.001**
Nutrition (High School)	694	22.94	389	23.66	.003*
Nutrition (Middle School)	217	20.92	149	20.53	.413
Nutrition (Grades 3 to 5)	441	22.00	364	22.39	.184
Parenting	498	91.45	178	97.00	*p<.0*01**
Science (HS and MS)	112	65.71	35	65.17	.829
Technology (Adults)	58	26.43	39	26.92	.653
Technology (HS and MS)	57	32.26	1	35.00	-
Workforce Preparation (HS)	227	44.43	160	46.46	.004*
Workforce Preparation (MS)	859	26.05	483	26.44	.077
Physical Activity (Grades 3 to 5)	114	26.18	71	25.70	.53
Technology (Grades 3 to 5)	488	5.36	438	5.47	.007*
Science (Grades 3 to 5)	431	10.94	311	11.33	.018*

Note: Only one participant completed the *Technology (HS and MS)* scale post-survey, therefore analysis was not conducted. * Indicates difference is significant at *p<.05 level; ** indicates difference is significant at p<.001* level.

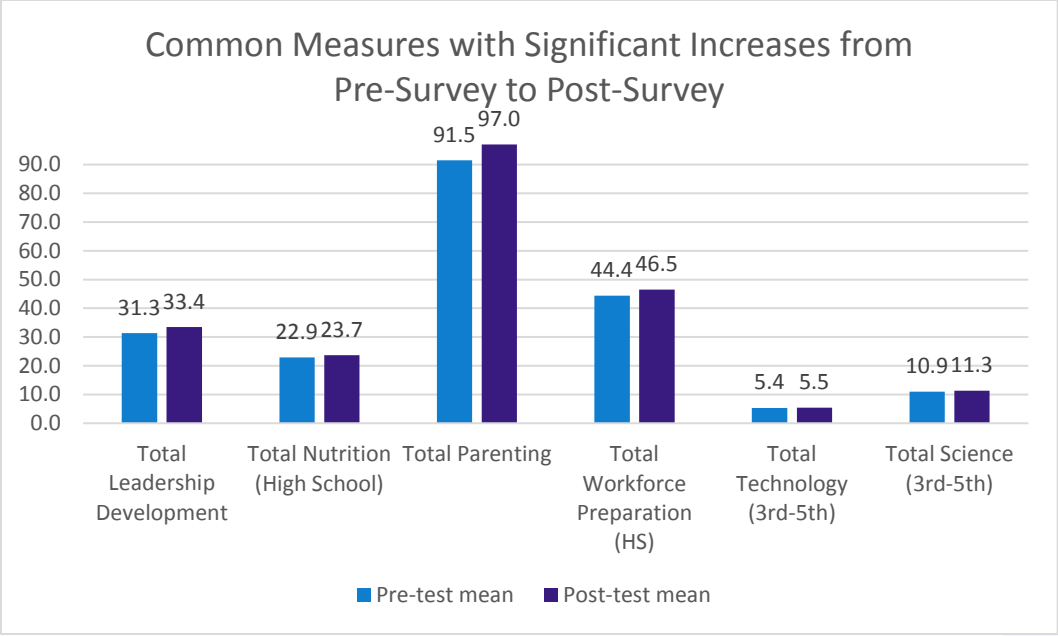

Figure 25. Significant increases from pre-survey to post survey were found on six common measures related to short-term outcomes.

Data Collection Conclusion

CYFAR data collection is increasing over time, with significant positive change beginning in 2014, when all SCPs were required to collect and submit common measure data. With this increase, though, has come a decrease in the proportion of youth (particularly in grades three through five) who are completing program quality and core competency measures. These measures are optional for youth in grades three through five; however, since that group of youth makes up about a third of all data collected, further investigation into this decrease may help to bolster overall data collection.

With regard to common measures of short-term outcomes, 89% of youth have completed at least one measure. Of the 11% of youth who have not completed at least one common measure, 80% are below grade three, for whom there are no common measures available. The consistently most-used common measures of short-term outcomes for youth are measures of workforce preparation and nutrition. For adults, parenting is an often-used common measure, with 58% of adults completing this measure.

Even though there has been a dramatic increase in the number of SCPs collecting data, there is still room for growth. With CYFAR serving approximately 10,000 individuals per year, SCPs are currently submitting data for at most 20% of participants.

This analysis demonstrates that there is room for greater quality in the data SCPs collect. Based upon the patterns of data collection over the past five years and analyses of those data, there are some strategies and adaptations that could decrease the amount of time and effort necessary for quality data collection and improve the overall quantity and quality of data collected.

One of the most effective ways of measuring the impact CYFAR programs have on youth is to measure the extent to which scores on the scales change from pre-survey to post-survey. In order to do so, the CYFAR PDTA Center established a naming convention for the development of unique identification numbers and emphasizes the need for pre-survey and post-survey data collection. When such data are available, one can examine the ways participants change during their time in a CYFAR SCP. This analysis would provide an accurate and powerful measure of change from pre-survey to post-survey.

Overall analyses of existing data demonstrate that there are significant relationships between program quality and positive youth outcomes. Youth who rate program quality higher tend to score higher on core competencies and certain common measures of short-term outcomes. There are also some significant differences in outcomes between pre-survey and post-survey groups. For core competencies, these changes are variable in that the post-survey group is lower on some outcomes (i.e., social conscience and personal values) than the pre-survey group, but higher on other outcomes (i.e., critical thinking and decision-making). With regard to common measures of short-term outcomes, youth in the post-survey group scored higher on leadership development and workforce preparation than youth in the pre-survey group. High school participants in the post-survey group also scored higher on nutrition than those in the pre-survey group; however, this difference did not exist for youth in grades three through middle school.

Conclusion

This report used cumulative cross-site data collected by CYFAR SCPs from 2011 through 2016 to provide an update regarding the progress of data collection efforts and evaluation of the relationship between CYFAR SCP programming and youth outcomes. A large programming culture-shift began in 2011 with pilot testing of the cross-site evaluation plan, and continued in 2014 when data collection became mandatory. In five years, programs have moved from collecting no cross-site evaluation data to a substantial number of sites regularly collecting data. This is a noteworthy change in a small period of time resulting in a continually growing amount of cross-site evaluation data. The ability to continue to evaluate CYFAR SCP participant outcomes will continue to depend on SCP adherence to the cross-site evaluation design.

Recommendations

There are adaptations that can be made to the current data collection system in order to assist SCPs and further adherence to data-collection conventions, enhancing both the quantity and the quality of data collected. There are two primary paths to improvement in data collection: increased fidelity to the original cross-site data collection design and improvements to the cross-site data collection design.

Continue technical assistance efforts. Currently, the CYFAR PDTA Center provides technical assistance for SCPs using the Survey Builder tool. This technical assistance enables programs to collect the highest-quality data possible. As a part of this on-going assistance, the CYFAR PDTA Center should inspect the Survey Builder tool and identify technical changes that can enhance the usability and efficacy of the tool. Within this effort, the CYFAR PDTA team should continue to be in close contact with SCPs to ensure awareness of the difficulties programs face in the Survey Builder system.

Add demographics to post-surveys. During the current data analysis, it became clear that having no demographic information available for post-surveys created difficulties at times. It reduced the ability to

analyze demographic factors affecting post-survey results, and to see if demographic differences in outcome measures were reduced at the time of the post-survey. Furthermore, it seemed that the lack of demographics on the post-survey led some programs to use the Survey Builder tool differently than was intended. Adding collection of demographics to the post-survey should: improve the ability to confirm a match between participant's pre-survey and post-survey, increase the breadth of analyses of post-survey data, and collect demographics for participants who enter the program after the pre-survey data has been collected.

Redesign questions regarding participation level. The measure of participation level obtained varied responses that made analysis of the data very difficult and conclusions regarding links between participation level and outcomes impossible. These questions should be reviewed and replaced with questions that will enable participants to provide clearer information, thus allowing for analysis of these data.

Recommendations for reducing the survey size and requirements. The CYFAR PDTA team has received feedback regarding the length of the current surveys. It is possible that if surveys were shortened, SCPs would be more likely to collect data and participants would be more likely to provide more accurate data. In order to simplify the data collection requirements, two major changes to the CYFAR common measures are recommended.

First, though program quality is currently measured at pre- and post-survey, participants lack knowledge at the pre-survey time to respond to questions about program quality. Thus, it is recommended that program quality only be collected during the post-survey, at which point participants will have a better familiarity with and understanding of the program environment. This will allow for higher-quality data regarding program quality.

Furthermore, it would be reasonable, considering the high correlations among the different measures to reduce the number of measured core competencies from five to two. This will decrease the length of surveys significantly without much loss of meaningful data. To this end, it is recommended that the decision-making and critical thinking be the only core competencies measured in the future.

It is the hope that these two changes will reduce the time and effort SCPs spend on data collection and measurement. In turn, this will likely increase the proportion of SCPs collecting data and increase the quality of those data.

Retrospective pre-test. Sometimes a retrospective design is used to evaluate change over time. A retrospective design would include collection of data only at the end of a program. With this design, participants would be asked to indicate where they are on a measure at the end of the program as well as indicating where they believe they were on the same measure at the start of the program. This eliminates the collection of data at multiple time points and eliminates dependency on analysts' ability to match pairs of pre- and post-surveys. This design can be an effective way of evaluating change over time, though there are some concerns, because it depends on participants' memory and accurate judgment of where they would have been at the start of the program. It would be useful to evaluate the efficacy of a retrospective design in the population of CYFAR participants in order to judge whether this could be a useful way to collect necessary data in the future.

Future Directions

ERIKA. ERIKA (Electronic Response to Information, Knowledge, and Action) is an online, illustration-based evaluation tool developed for children between the ages of six and twelve, with particular focus on children and youth whose reading skills may not be sufficient to participate in current surveys. The tool is currently being designed for the measures of core competencies. It is hoped that ERIKA will increase the number of children and youth in grades one through five who complete measures of core competencies.

C-NAT. The CYFAR-Nutritional Assessment Tool (C-NAT) is an illustration-based tool also designed for youth ages six through twelve. This tool specifically aims to simplify the collection of the nutritional assessment data for children and youth in grades one through five who participate in CYFAR SCPs.

Evaluation Institute. The CYFAR PDTA team is currently developing the Evaluation Institute to increase evaluation capacity among CYFAR grantees. The Evaluation Institute will engage program evaluators in a series of interactive webinars and other activities that will help them learn and teach SCP teams about the structure, function, rationale, and procedures of cross-site data collection. The ultimate goal of the Evaluation Institute is to improve both the quality and quantity of the evaluation data collected by SCPs. The Evaluation Institute is being piloted with a select group of SCP evaluators and is expected to launch summer of 2017.

eXcel. High program quality is associated with better youth outcomes. As such, the CYFAR PDTA center has created Excellence in Youth Programming (eXcel). This is a virtual coaching process that supports youth development programs in achieving and maintaining quality. eXcel offers a comprehensive process for assessing and understanding key aspects of program quality by using observational data to identify the program's strengths and challenges and to design and implement a program development plan to address those challenges. CYFAR coaches will use eXcel tool to provide virtual technical assistance and training to improve programs and in turn improve the experience and impact for youth in CYFAR programs.

CYFAR SCPs are increasing their collection of data over time. With the implementation of these recommendations, we hope to enable SCPs to collect more high-quality data. This will allow for further examination of the possible influence of CYFAR on positive youth outcomes.