



Measuring Youth Development: How Out-of-School Time Programs Collect and Use Data

Hannah Lantos, Zakia Redd, Jessica Warren, Michael Bradley,
and Sham Habteselasse

Contents

Acknowledgements.....	ii
Executive Summary.....	1
Introduction.....	5
Chapter 1: Methods.....	7
Chapter 2: Outcome Measurements.....	10
Chapter 3: Factors Programs Consider When Deciding What Outcomes To Measure.....	16
Chapter 4: Measurement Challenges.....	20
Chapter 5: Key Findings And Recommendations.....	24
Appendices.....	24
Appendix I. Additional References.....	29
Appendix II. Bibliography For Ost Measurement Literature Scan.....	31
Appendix III. Expert Informants.....	49
Appendix IV. Programs That Participated In The Study.....	50
Appendix V. Data Collection Protocols.....	52
Appendix VI. Survey.....	58
References.....	62

Acknowledgements

Numerous people helped develop this report. We would first like to recognize and thank the individuals from the 38 programs who participated in the study and shared their experiences, insights, and data collection tools through interviews and surveys. We also thank the 23 experts for providing recommendations for programs to include in this review.

In addition, we thank Mark Waits and Brent Franklin of Child Trends for their careful reading and editing of this report and helpful suggestions. We also thank our senior reviewers, Kristin Anderson Moore and Emilie Phillips Smith, for their thoughtful critique and extensive comments.

Finally, we express our gratitude to The Wallace Foundation for commissioning and funding the research on which this report is based. We especially thank Bronwyn Bevan for her ongoing guidance, substantive feedback, and support. We also thank Gigi Antoni, other members of the Youth Development team, and Pam Mendels and Daniel Browne of Wallace's editorial team for their insights, recommendations, and review.

Suggested citation: Lantos, H., Redd., Z., Warren, J., Bradley, M, & Habteselasse, S. (2024). Measuring youth development: How out-of-school time programs collect and use data. Child Trends. DOI: 10.56417/3703u3452e

Executive Summary

Out-of-school-time (OST) programs and their funders rely on sound data to make decisions about everything from professional development and student recruitment to the selection of activities to offer students. Programs operate at a range of times (before and after school, weekends, summer) and in a variety of locations (e.g., schools, community-based organizations, city parks and recreation centers), are run by a variety of entities (e.g., government agencies, private community organizations), and receive funding from a variety of sources (e.g., government, philanthropy)—each of which may be interested in a different set of data and come with its own reporting requirements. This means there is a great deal of variation in the types of data programs collect.

In 2019, The Wallace Foundation (Wallace) commissioned Child Trends to conduct a study of the kinds of youth outcomes OST programs are interested in measuring, the tools they use to measure those outcomes, and the challenges they experience in doing so. The study included a literature scan and interviews with leaders and staff members at 28 OST programs. Twelve of the 28 also completed surveys; a separate group of 10 provided information by survey only. The study expands on past research by a) focusing on programs that work in specific content areas (e.g., the arts, civic engagement and social justice, career and workforce development) and b) covering both *quantitative* approaches (i.e., tracking numerical data) and *qualitative* approaches (gathering descriptive information through surveys, interviews, etc.) to data collection.

While its findings apply to OST programs in general, the study focused on particular types of programs (i.e., afterschool, summer, online) and particular content areas, as well as programs that serve school-age children and adolescents from marginalized communities, those that support students' social and emotional learning (SEL), those that serve systems-involved youth, and those that focus on promoting equity—for example, by training staff to recognize and overcome personal biases or by recruiting and retaining leaders and staff who reflect the diversity of the participants served.

Key findings

What outcomes do programs measure?

- Programs in the study measured outcomes that were closely related to the content they delivered (e.g., the arts, career and workforce readiness, civic engagement and social justice). They also generally measured SEL outcomes and other outcomes required by funders.
- Programs thought critically about how to measure progress toward achieving equitable outcomes for the youth they serve. Approaches included disaggregating data by race and other demographic variables and tracking the development of equity-related knowledge and skills.
- Programs used a number of criteria to decide what outcomes to track, including whether a given outcome was consistent with the program's logic model or theory of change, usability and shareability of the data, the effort and capacity required to collect the data, availability of valid and reliable measures, youth interest, and whether measuring a given outcome would promote equity.
- In addition to outcomes, programs consistently measured outputs (i.e., steps that lead to desired outcomes), particularly program participation and quality. This finding aligns with our literature scan that revealed that public and private funders have invested heavily in quality assessment tools and participation tracking systems. Comparatively few programs used qualitative methods to understand the factors that contribute to program participation and quality.

What methods do programs use to measure progress?

- Programs reported using a variety of traditional methods to measure outcomes, including *quantitative* methods—like administering surveys and questionnaires, conducting formal assessments, and tracking attendance rates—and *qualitative* methods such as conducting interviews and focus groups with young people. Their use of quantitative methods was well-documented, but their use of qualitative methods less so.
- Programs also used nontraditional methods, such as regular check-ins with participants and reviewing participant journals and portfolios, along with creative, informal ways to track SEL outcomes, including games and award systems. As with their use of qualitative measures, use of these nontraditional methods was not well-documented.

What measurement challenges do programs face?

- The programs we interviewed did not consistently identify specific outcomes of interest that they were unable to measure, although some said they lacked the tools they needed to track equity- or SEL-related outcomes or program quality.
- Programs pointed to broader types of outcomes that were challenging to measure, including longer-term outcomes like college matriculation, career attainment, and participation in civic life; behavior change (e.g., whether participants in a civics program exercise their right to vote); school outcomes like test scores and grades that require a data-sharing agreement with the district; and the relationship between improvements in staff members skills and knowledge and youth outcomes.
- Programs reported that the process of collecting data could be burdensome for both participants, who have their fill of testing at school, and staff, who in many cases do not have the training or time to do the work. Some pointed specifically to the burden of fulfilling the reporting requirements of various funders, which often involved recording duplicative information using multiple tools and databases.

Recommendations for OST programs

Select a small number of high-priority outcomes to measure.

Programs often cast a wide net for data and end up not using the data they collect, which can be frustrating for staff and participants. To avoid this, programs should focus on data they will be able to use and share with participants, families, and the broader community.

Ensure that the selected outcomes fit into the program's logic model or theory of change.

Many programs reported they did not use their logic model or theory of change to decide what outcomes to measure or guide their data collection efforts. Programs should review the outcomes they target for consistency with their logic model or theory of change and, if necessary, make changes to one or the other.

Document the type, source, and purpose of each piece of data collected.

Programs were more likely to keep a record of the data they collected using traditional, quantitative methods like surveys or formal assessments than data collected using qualitative or nontraditional methods like check-ins with participants or portfolios of their work. Programs should thoroughly document all the different ways they collect data to ensure that all data get used and to better understand what each method contributes to their work.

Make room for youth perspectives in decision making about what outcomes to measure and how.

It is important that programs find out what young people want to get out of their participation (e.g., to learn a particular skill or to get a job in a particular field) and their progress toward those goals. Programs can solicit input from participants not only in deciding what outcomes to measure but also in assessing how well the data collection process is working and in determining the meaning of data collected.

Incorporate measures of diversity, equity, and inclusion into data collection efforts.

Programs interested in tracking their progress on issues of equity should consider disaggregating data by race or other demographic variables to identify disparities in participant outcomes; incorporating outcomes specifically related to equity (e.g., the development of civic knowledge and social justice skills) into their logic models or theories of change; revising their curricula to promote those outcomes; and selecting qualitative and, if possible and appropriate, quantitative methods of measuring those outcomes.

Assess the capacity and effort required to collect data and take steps to minimize burnout.

Programs should expect finding the right approach to data collection to be an ongoing process of trial and error. One way to reduce the burden on staff is to schedule data collection for a few specific points in the calendar when data will be most relevant rather than trying to collect it continuously.

Recommendations for funders

Connect programs to additional training in key aspects of data collection and use.

Many staff members who oversee their programs' data efforts are trained in youth development and are learning to work with data on the job. Interviewees spoke of the need for more training in assessing which measurement tools are valid and reliable; developing new databases or navigating the systems they have; and using data to support program improvement.

Create networks of programs to provide informal technical assistance.

One program secured a grant to create a network of OST programs that could support and learn from one another. Through this network, the program learned how other programs have solved common problems and worked with those programs' technical assistance specialists to troubleshoot its own particular challenges. This program noted it would also benefit from participation in networks focused on developing specific skills (e.g., collecting longitudinal data, identifying outcomes of interest and ways to measure them).

Provide funding for additional staffing.

Tight budgets make it difficult for programs to hire dedicated staff for data collection and analysis, and the existing staff are focused on providing the services youth need and don't have sufficient time to support data-related tasks. Additional resources would allow programs either to hire staff whose primary responsibility is data or reduce staff-to-participant ratios so all relevant staff members have more time for data activities.

Partnerships between practitioners and researchers can be mutually beneficial. Researchers can provide programs with technical assistance in identifying what outcomes to measure, selecting valid and reliable measurement tools, and analyzing the data they collect, while programs can provide researchers with real-world data to analyze and an opportunity to field-test the tools they develop.

Make reporting requirements as flexible as possible.

Almost every program in our study mentioned that meeting the requirements of their various funders was a time- and resource-intensive process that involved recording information that was duplicative or not meaningful to them, using multiple tools and databases. Funders should be sensitive to the strain that data collection can place on programs and be as flexible as possible in terms of how and in what format grantees demonstrate progress toward outcomes of interest.

Support digital data collection.

As many programs moved online during the COVID-19 pandemic so did their data collection efforts. Even as they return to serving young people in person, programs could benefit from additional funding for the development of mobile data collection processes and training to do this work.

Recommendations for further research

Expand the existing study to include more programs and voices.

The current study could be expanded by including a) more programs in each of the categories we prioritized; b) programs in additional categories (e.g., STEM, physical fitness, identity development) to increase the generalizability of the findings; and c) perspectives from different types of people involved with OST programs, particularly youth participants.

Further explore how programs make decisions about their data efforts.

Data collection and use require program leaders to make difficult tradeoffs in terms of money and staff time. To help programs navigate these tricky decisions, future research could focus on the following:

- How do programs identify outcomes meaningful to them and prioritize those outcomes?
- How do they choose the most appropriate measurement tools for their needs?
- How do they balance spending on improved data collection with spending on programming?
- What types of support for data collection and use (e.g., training, financial, staffing) do they find most helpful?

Highlight innovative solutions to common problems.

Future research could identify how exemplary programs tackle common data-related problems. Topics could include forming successful data-sharing partnerships, minimizing the burden of data collection on staff, building staff capacity to collect and analyze data, and securing funding for data efforts.

Investigate strategies for collecting longitudinal data.

Programs expressed interest in tracking long-term participant outcomes but had difficulty doing so. Many programs specifically mentioned the challenge of collecting data from alumni who may no longer be connected to the program. Future research could identify strategies for effectively measuring long-term outcomes by conducting a formal review of best practices in the field and gathering qualitative evidence from programs that have done it successfully.

Catalog existing measurement tools and develop and test new tools as needed.

Through our literature scan, we identified a robust set of tools for measuring program quality and SEL. Yet some programs were unaware of them. Researchers could help address this disconnect by cataloging the

tools already in use in the field and assessing their validity and reliability. The literature scan also revealed a lack of tools relevant to content-specific, equity-related, and other types of outcomes. Researchers could develop new tools to fill this gap. Funders could support this work by putting out a call for research in field publications; organizing panels or conference sessions focused on this topic; or forming a network of practitioners, funders, researchers, and content experts to identify and share promising approaches. Funders could also spearhead efforts to get new and existing tools into the hands of programs.

Introduction

“What gets measured gets done,” the saying goes. The idea that collecting and using data can play a powerful role in shaping an organization’s work certainly applies to out-of-school-time (OST) programs. OST programs and their funders rely on sound information to make decisions about everything from professional development and student recruitment to selecting activities to offer students.

The OST field, which includes afterschool, summer, weekend, and before school programs, is broad and diverse. Programs operate in a variety of locations (e.g., schools, community-based organizations, city parks and recreation centers), are run by a variety of entities (e.g., government agencies, private community organizations), and receive funding from a variety of sources (e.g., government, philanthropy), each of which may be interested in a different set of outcomes and come with its own reporting requirements. For example, the 21st Century Community Learning Centers program focuses primarily on academic outcomes, while Child Care and Development Block Grant funders may be more interested in understanding the extent to which programs help meet the needs of working parents. The breadth and diversity of OST programs, funders’ requirements, and characteristics result in significant variation in the types of data programs collect.

Historically, there has been a mismatch between the type of benefits OST programs typically confer and the type of data they collect. The field has long emphasized the social and emotional development of young people, and yet programs primarily tracked outcomes related to academic achievement. Recognizing this mismatch, The Wallace Foundation, a long-standing funder of OST, along with other funders, has invested in developing, testing, and cataloguing measures of social and emotional outcomes for OST programs.

In 2019, Wallace commissioned Child Trends to study the kinds of youth outcomes OST programs are interested in measuring, the tools they use to measure those outcomes of interest, and the challenges they experience in doing so. The study included a literature scan and interviews and surveys of OST program leaders and staff members. It expands on past research by a) focusing on programs that work in specific content areas (e.g., the arts, civic engagement and social justice, career and workforce development) and b) covering both *quantitative* approaches (i.e., tracking numerical data) and *qualitative* approaches (gathering descriptive information through surveys, interviews, etc.) to data collection. It is common for OST programs to use qualitative methods but less common for them to use such methods in formal program evaluations.

While its findings apply to OST programs in general, the study focused on particular types of programs (i.e., afterschool, summer, online) and particular content areas, as well as programs that serve school-age children and adolescents from marginalized communities, those that support students’ social and emotional learning, those that serve systems-involved youth, and those that focus on promoting equity, for example by training staff to recognize and overcome personal bias or by recruiting and retaining leaders and staff who reflect the diversity of the participants served.

This report is organized into the following chapters:

- Chapter 1 provides an overview of the methods used to collect information for this study.
- Chapter 2 focuses on findings related to outcomes valued and measured by the OST programs interviewed for the study and the measurement techniques used to gather information on outcomes.

- Chapter 3 highlights factors programs considered when measuring outcomes.
- Chapter 4 focuses on challenges experienced by programs as they collected or tried to collect data on outcomes of interest.
- Chapter 5 includes key findings from the study and recommendations for the field.

Appendices are included at the end of the report to provide additional contextual information on our approach and findings.

Chapter 1: Methods

Research questions

The study was designed to answer the following three research questions:

1. What is the **range of intended outcomes** for children and youth attending OST programs, including, but not limited to, programs that focus on the arts, civic engagement and social justice, career and workforce development, and general or other services?
2. What are the **different measurement approaches and instruments** used to document OST program outcomes, including formative measures (i.e., assessment that occurs during the program) and summative measures (i.e., assessment that takes place at the end of the program)?
3. What are the **gaps in and barriers to documentation and measurement** where intended outcomes are either not measured or are not articulated because of a lack of a measurement strategy, skills, resources, or instruments?

The first step in answering these questions was to conduct a scan of the literature for studies of OST programs focused on measuring outcomes in the areas listed in Table 1 below. Studies were considered “relevant” if they focused on measuring outcomes. Studies were categorized as “not relevant” if they were about program quality or implementation or were descriptive in nature.

Table 1. Studies of OST Programs Focused on Measuring Outcomes, by Outcome Type

The 283 relevant studies included outcome measures in the following areas:

	# of Relevant Studies
Social and Emotional Learning	93
Academics	74
STEM	60
Arts	23
Identity and equity	22
Civic engagement	11
<i>Unclear from abstract</i>	15

Nearly a decade has passed since the last comprehensive reviews focused on measurement tools and strategies and data use in OST programs were published. (Wilson-Ahlstrom et al., 2014; Hair et al., 2001; Bronte-Tinkew, Moore, and Shwalb, 2006). More recently, there have been several reviews of findings from OST programs (McCombs, Whitaker, & Yoo, 2017; Sloan McCombs et al., 2019; McCombs et al., 2017; Lantos et al., 2021). This study did not focus on the types of outcomes OST programs are likely to produce but rather on how programs identify and track those outcomes.

Sampling approach

Following the literature review, we set out to identify a sample of OST programs with the intention of conducting surveys and in-depth interviews with staff members to understand the outcomes they prioritize, how they measure those outcomes, the barriers to successful measurement, and the complex decisions they

must make regarding the selection and measurement of outcomes. The purpose of the brief, initial survey was to help the interviewers gather background information (including the program’s theory of change or logic model and data collection tools), so they could probe with deeper, more complex questions during the interview.

We decided to target OST programs identified as strong by experts in the field. We hypothesized that strong programs would be more ready for evaluation and have greater capacity for data collection than the average program. Narrowing our focus to a subset of all OST programs would also help us minimize irrelevant variables and reach “saturation,” the point at which we had conducted enough interviews to confidently identify common themes.

In order to explore how programs measure equity and any barriers to doing so, we used a purposive sample (i.e., a sample in which participants are selected based on desired characteristics), focused on programs serving diverse populations, communities of color, and justice-involved youth.

Child Trends asked 35 experts in the OST field—including leaders from national policy and advocacy organizations, national youth-serving organizations, researchers, and funders—to recommend two to five strong, high-quality programs to be part of this sample. (See Appendix III. for the list of the experts who provided program recommendations). Experts possessed broad knowledge of the OST field or specific knowledge about a segment of the field, such as summer programs, arts programs, or programs that serve youth engaged in the juvenile justice system. Outreach to the experts consisted of an email that briefly described the study, our purpose for reaching out to them, how they were referred to us, and the types of programs in which we were interested. We offered to have a videoconference, if necessary, to gather recommendations, but most experts provided recommendations by email. We felt it was important to allow each expert to define program quality independently. Thus, we did not define high quality for them in our outreach email and instead asked for “recommendations of strong programs to interview.”

Programs in the study

The 20 experts who responded to our request recommended 105 programs that they considered strong. We selected 45 programs out of the 105 recommended to contact for interviews using videoconferencing software. To select this first set of programs, we prioritized four criteria: 1) programs recommended by multiple people, 2) program types that Wallace identified as priorities (afterschool, summer, online, programs that serve diverse populations or focus on equity), 3) program types for which we had fewer recommendations from our experts overall (e.g., online), and 4) programs that did not have recent, rigorous evaluations. If two programs shared these criteria, we prioritized the ones that would add geographic variation to the sample. When responses from this first set of programs slowed, we reached out to 13 programs that were either online programs or career readiness programs, as these were the two program types for which responses were still low (only one or two responses each). In the end, 28 programs participated in interviews, and an additional 10 provided information by survey only. We also received fully or mostly completed surveys from 12 of the 28 interviewed programs. Our final distribution of interviewed programs can be found in Table 2.

Table 2. Number of Programs Interviewed by Category

Focus Area	Total
Arts	7
Career & workforce	5
Civic engagement & social justice	7
Social and emotional learning	5
Programs with specific timing, location, or target population	
Online	3
Summer camps	3
Serves youth engaged in the juvenile justice system	5
Serves young school-aged children	11

Note: This table includes only the programs we interviewed, not the programs that completed surveys but were not interviewed. Also, the total here adds up to more than 28 because programs sometimes fell into multiple categories (e.g., social justice and arts). Some of the programs listed in the bottom half of the table (those with specific timing, location) overlap with the focus areas in the top half (e.g., online and social and emotional learning or juvenile justice and art).

Chapter 2: Outcome Measurements

Valued and measured outcomes

In this section, we discuss the outcomes on which programs focused. Specifically, by outcomes we mean the knowledge, attitudes, behaviors, or skills that programs want the young people they serve to develop. Many programs think about short-term outcomes (e.g., knowledge gained) and long-term outcomes (e.g., behavioral change) but measure only short-term outcomes.

Before summarizing the outcomes that programs reported measuring, we present definitions of key measurement-related terms used throughout the rest of this report, using “well-being” as an example of an intended outcomes.:

Table 3. Definitions of Key Terms (Using “well-being” as an example)

Domain	The broadest conceptual category of well-being. For youth-serving programs, domains may include educational well-being, social and emotional well-being, economic well-being, and physical health, among others. Within each domain, there may be several constructs (defined below).
Construct	A conceptual category within the specific domain of well-being. For instance, a program may be interested in assessing educational constructs of academic performance, school engagement, or educational attainment or social constructs such as sense of belonging or collaboration.
Measure	The actual metric used to assess participant well-being. These may involve specific tools and methods for data collection and reporting, such as surveys, scales, observational codes, administrative records, etc. There may be several ways to measure a construct of interest, such as academic performance or sense of belonging, depending on the age group of the students in the program, among other factors. For younger children, a measure of belonging might entail a parent or teacher assessment. For older children, programs might use a self-report survey.
Output	Measure of program participation and services offered, such as the number of students served, the amount and type of sessions or supports provided, and the number and type of projects participants complete.
Outcome	Measure of the changes experienced by the target population as a result of participating in a program. Outcomes include changes in knowledge, attitudes, behaviors, and skills and may occur over a short or a long period.
Mediator	An intermediate output or outcome that influences whether a program achieves a targeted outcome. For example, frequency of participation in a program, duration of instruction, or program quality may serve as mediators for whether a program achieves a targeted outcome like improved social and emotional skills.
Moderator	A specific condition or characteristic of the target population that leads to variation in outcomes. For example, boys participating in a program may benefit more or less, or differently, than girls.

The term “measures” may refer to outputs, outcomes, mediators, and moderators. Each of these is important in informing programs’ efforts to reach their long-term goals and improve program quality.

Content-specific measures

The most easily identifiable measures were specific to the content of the program (e.g., arts, career and workforce readiness, civic engagement and social justice). See Table 4 for examples. Table 4 also includes commonly reported measures, such as social and emotional learning (SEL) and academics, that were not specific to a particular content area. Some programs identified other measures, such as physical activity, as a focus. Most programs spoke of these measures as either short- or long-term outcomes, though some could be thought of as mediators (e.g., improved math grades or test scores being a step toward graduation).

Table 4. Examples of Program-Specific Measures

Focus Area	Examples of Measures
Arts Programs	Completion of a performance; creative thinking skills; artistic ability (dance, music, mural painting, etc.); public speaking skills; self-expression; community engagement; knowledge of how to use the arts to engage others
Career & Workforce Programs	Employment; earnings; knowledge of career options; completion of sample resumes or cover letters; positive reviews from supervisors; completion of an internship; credentials earned; job readiness skills (professional communication, timeliness, job-specific technical skills, etc.); development of a professional network
Civic Engagement & Social Justice Programs	Knowledge about government systems and processes; positive attitude toward civic engagement; intention to participate in civic life (e.g., attend a city council meeting or email an elected representative); ability to perform root cause analyses (defined below); knowledge of political and social systems; ability to listen to other points of view
Online Programs <i>(Note: Online programs in our sample were mostly STEM-focused.)</i>	Technical skill mastery (robotics, computer science, gaming, e-sports, etc.); teamwork; age-appropriate math skills; digital citizenship (defined below)
SEL Outcomes	5Cs +1 (competence, confidence, connection, character, caring, and contribution); self-esteem; self-efficacy; cooperation; teamwork; empathy; problem solving; decision-making skills; resilience
Academic Outcomes	Graduation; on-time advancement to next grade; college enrollment; passion for learning; school engagement; ability to identify universal themes in literature and theater; English and math skills
Other Outcomes	Life skills, physical activity, reduced recidivism, equity-focused outcomes (see more below)

Below we describe two categories of measures—SEL and equity—in more detail.

Measurement of social and emotional learning (SEL)

Our interviews with OST programs clearly supported that SEL has been a growing priority in the OST field over the past decade. SEL outcomes (e.g., relationship-building skills, emotion management, communication skills, empathy) were a point of emphasis for the programs in our sample, regardless of program content. Arts, career readiness, and even online programs focused on video gaming incorporated SEL. For example, a theater program might stress the importance of teamwork, while a civics program might make a point of building students' empathy. One interviewee noted that, while civics and SEL are closely linked, it is easier to measure—and get support for measuring—SEL:

"[SEL] is naturally embedded in what [...] civic development goals are anyway, like developing critical thinking... From a measurement standpoint, the field of SEL is much more developed and has a lot more standardized [measures] we can pull from and also has much more interest and investment from funders."

Some programs explicitly measured SEL outcomes by weaving specific SEL measures (either validated measures from other surveys or questions they developed on their own) into the other surveys or assessments they already conducted.

Several programs that focus on younger children did identify SEL as their primary content area. These programs often had curricula geared toward developing these skills, and measures were explicitly linked to those curricula. For example, [Wings for Kids](#), an SEL program that works with elementary school-aged children in three southern states, focuses on and measures five core SEL competencies: self-awareness, self-management, responsible decision making, social awareness, and relationship skills. The program uses the mini-Devereux Student Strengths Assessment ([DESSA mini](#)) twice a year to measure these five competencies. (For a short time, they also collected weekly assessments of students; however, this was burdensome for staff and ultimately resulted in poor quality data.) The DESSA is a 72-question screen that takes 5-10 minutes to complete and measures eight SEL competencies, while the DESSA mini consists of eight questions and takes one minute to complete (see [here](#) for more information about both instruments). [Wings for Kids](#) also collected some qualitative data, primarily through interviews and focus groups, to better understand how program quality and implementation are related to outcomes. Multiple programs said they use a different measurement tool, [Hello Insight](#), to collect data on SEL.

Measurement of diversity, equity, and inclusion

It was clear from our interviews that the rise of the racial justice movement in 2020 prompted some programs to think critically about their curricula and the outcomes they do or do not measure. Whether this reflection led to changes in measurement strategies varied from program to program. Many interviewees spoke of the need for programmatic changes related to diversity, equity, and inclusion (DEI) and said that those changes would help them determine what to measure. Following are three other themes about how programs measure DEI that arose from our interviews:

Disaggregating data by race/ethnicity

One way programs assess their progress on issues of equity is by determining whether participants from different backgrounds experience equitable outcomes. Specifically, by disaggregating data by race and ethnicity or other demographic variables (age, gender, neighborhood, socioeconomic status, etc.), programs can begin to understand whether their program is helping to reduce disparities, having no impact, or even exacerbating them. Disaggregating data is an analytic strategy, not a measurement strategy; however, it is related to measurement because it necessitates collecting data on race and ethnicity or other demographic variables of interest. Some programs did this well, while other programs noted that most participants were youth of color and thus disaggregating by race was not feasible. These programs sometimes identified other factors that they could disaggregate such as income, neighborhood, and length of time in the program. These

factors could serve as proxies for risk or vulnerability in the same way that other programs sometimes used race. Many programs did not yet have a clear process for responding to lingering or worsening disparities they identified.

Tracking development of equity-related knowledge and skills

Some programs, particularly arts, civic engagement, juvenile justice, and camp or recreation programs, have the development of civic knowledge or social justice skills among their intended school outcomes. These skills involve learning how to discuss racism, power, inequity, and their personal experiences with these complex social dynamics and, in some cases, engaging with members of the community to address the effects of racism. The programs sometimes assess these outcomes using traditional methods such as surveys but more often using nontraditional methods like discussions, the art students produce, or the local action campaigns they carry out. In addition, equity is often embedded in the content of programs that do not make it an explicit focus of their work, for example by choosing equity-related topics for a theater production, facilitated group sessions, or local campaign work. Just as a program that does not explicitly focus on SEL might track SEL outcomes informally, programs that do not explicitly focus on equity might ask staff members if they feel there are any equity issues to discuss or whether disparities appear to be arising.

Revising curricula and/or outcomes measures based on equity reviews or perspectives

In redesigning their curricula in 2020 and 2021 to better address racial equity issues, some programs more clearly described how they would track students' knowledge of these issues. For example, a civics education program spoke about how an exercise called root cause analysis, which encourages students to think backwards from an issue to its underlying causes, had always been a key part of its curriculum but would now explicitly include discussion of inequities in power and resources. For example, students might identify differences in high school graduation rates for different groups of people and then work out an explanation for this outcome. Teachers would guide those conversations so that power, inequity, resource allocation, and representation were explicitly included in their analyses. An interviewee described the new approach:

"We've been working with some DEI consultants, and we have some great people on staff also who are fully overhauling the curriculum... [so] a lot of the civics knowledge aspects and the civics skills aspects are more clearly grounded in students' learning around what systems of oppression are, how they function, how their identity and intersectional identity shape their experience in classes and communities. And then, hopefully through that, supporting a much more grounded root cause analysis and policy-action approach for our youth."

For this program, increased knowledge of how power and resources are used, and subsequent behavior change are long-term outcomes; whether discussion of these issues takes place is a mediator. In other words, the program's assumption is that students' knowledge about equity-related issues will increase if discussion of those issues is part of their root cause analyses.

Measurement tools and techniques

Programs reported that they use a variety of traditional and nontraditional methods to measure targeted program outcomes with varying degrees of success. Traditional methods of data collection include observations, surveys and questionnaires, interviews, focus groups, formal assessments, observation of participant performances, self-assessments, tracking attendance rates, and collecting school data. Nontraditional methods, including check-ins and reviewing participant journals and portfolios, often blend more seamlessly into programming than traditional methods because they are generally designed to be instructional tools first, evaluation tools second. In fact, while many programs have used check-ins, journaling, and portfolios for years, in some cases for quality improvement purposes, some still do not recognize that these methods can be used to measure student outcomes. Many stated that they think outcomes measurement needs to be more formal or quantitative.

Traditional methods

Surveys

Almost all programs reported that they use surveys to collect participant data. Programs used a variety of surveys, including validated tools (i.e., tools that have been developed by experts and shown to be effective in the field), and more informal surveys that they developed themselves. (Programs that reported developing their own survey explained that they did so because other surveys did not ask the questions they needed.) While most programs administered surveys at the end of the program or sporadically during the program, many administered a pre- and a post-survey at the beginning and end of the program.

For online programs, in which participants never meet at the same time, it may not be practical to administer a survey at a specific point in the program to the entire population. One online program used quick randomized pop-up surveys that participants, and often staff, completed during their online sessions. A few programs also administered a survey to parents.

Programs also used tools validated by universities and other organizations for participants to conduct self-assessments. These tools varied by program and, in some cases, changed from year to year.

Formal assessments of knowledge and skills

Some programs, including programs focused on STEM, SEL, and juvenile justice, used formal assessments to measure not only what and how well the participants learned but also their proficiency in applying what they learned. These formal assessments varied from program to program and included short quizzes, skills tests, and questionnaires.

Fine arts programs that focused on poetry, music, literature, and dance typically used performances at the end of the program—such as recitals or choreographed dances—to measure whether and how well participants learned relevant skills. In one theater program in Chicago, participants worked together to write their own play drawing on stories from their peers, communities, and their own lives. The final performance demonstrated the extent to which the participants had developed storytelling and other skills.

Observations

Just fewer than half of the programs in our study reported using observations by staff members to measure participant outcomes. In many cases, this method of measuring outcomes entailed observing participants as they worked on projects individually or in teams to gauge their progress toward a specific goal (e.g., developing interview skills) and including this information in reports throughout the program or at the end of the program. As noted above, arts programs in our study commonly observed performances to measure participants' progress in developing particular skills.

Interviews and focus groups

Several programs conducted interviews and focus groups to collect information from participants, staff, and parents that they then used for quality improvement purposes and for measuring youth outcomes. Some programs used interviews and focus groups to learn why youth might be struggling or to better understand how to support them. Additionally, a few programs, mostly arts programs, used interviews and focus groups to gather participants' own perceptions of what they had learned.

School data

Many OST programs focus on academic skills. Therefore, almost all the programs we spoke to wanted access to school data—administrative records, student grades, or teacher reports—that they could use for quality improvement and evaluation purposes. Very few, however, actually had access to this data from the schools or school district because it tends to be carefully protected and rarely shared. Programs that did have data use agreements with schools or school districts were able to use that data to track participants' progress in

school over the course of the program. A few programs that did not have such agreements asked students or parents to self-report grades or test scores.

Information programs tracked beyond outcomes

Many programs noted that they collected and measured information other than participant outcomes. For example, many programs reported tracking attendance, an output that is usually considered a mediator of behavioral outcomes like skill development, knowledge, or behavior change. As with other program outputs, attendance is easier to track than outcomes like increases in knowledge or changes in behavior. Some programs used attendance as a proxy for engagement (i.e., the degree of young people's participation in the program) and as part of their reporting to funders.

Many programs also reported measuring program quality using assessments such as the Weikart Youth Program Assessment (YPQA) or the After-School Quality (ASQ) Observation Tool.

Nontraditional methods

Informal measurement of SEL outcomes

Programs, particularly summer camps, found creative ways to track SEL outcomes. One was to use games or an award system. For example, [Fiver](#), a summer camp and school-year program, developed a system of “dog tags” that youth get for exemplifying specific competencies such as trustworthiness. The youth wear the dog tags proudly, even when they return home from camp or go to college. Giving awards for key program outcomes made it possible for Fiver to track participants' progress over the course of the summer without additional assessments or surveys and incentivized youth to make further gains. The program could also track the acquisition of dog tags on aggregate year over year, but it was unclear whether it did so.

Online programs also informally tracked SEL outcomes that were important to them. Staff and coaches at [Connected Camps](#), an online program that partners with schools to support students' participation in an e-sports league, met regularly to discuss students' progress toward SEL outcomes like teamwork, sportsmanship, commitment to racial and gender equity, sense of community, and digital citizenship (i.e., treating people respectfully online) and to identify ways to help students struggling in these areas. One staff member commented: “We have a big emphasis on healthy gaming, reducing toxicity...conflict solutions, collaboration...a lot of those ‘soft skills’ within the digital citizenship and SEL categories.” Connected Camps staff added that these skills are essential to STEM in particular and career readiness in general. They explained that many of the young people they serve were unable to represent their schools on a sports team, so e-sports offered a vital opportunity for building SEL skills.

Regular check-ins

Many programs reported using one-on-one or group check-ins with youth throughout the program to ask them how they were doing. Some programs scheduled check-ins, while others engaged with their participants informally. Many programs also used staff meetings or professional development sessions as informal opportunities for staff to assess progress toward SEL- and equity-related goals.

Journaling

The two programs that reported using journaling as a method of collecting data said participants were instructed to respond to prompts given by staff or to free-write. Staff members reviewed the journals throughout the program to gauge participants' learning.

Portfolios

Several programs, including arts, career and workforce readiness, and juvenile justice programs, used portfolios collecting participants' artwork, job records, or other relevant materials to measure outcomes.

Chapter 3: Factors Programs Consider When Deciding What Outcomes to Measure

This chapter focuses on how programs decide what outcomes they want to measure. These decisions are driven not only by the programs' theories of change (i.e., their understanding of how their work influences participants' knowledge acquisition, skill development, or behavior change) but also practical considerations like the cost of data collection. Program staff reported several essential criteria for deciding what outcomes to measure:

- Does the outcome fit into the program's logic model or theory of change?
- Is the data usable?
- Is it shareable?
- Does it incorporate youth voice?
- How much effort and capacity are required to collect the data?
- Are available measures valid and reliable?
- Does measuring the outcome promote equity?

Logic models or theories of change

Many programs reported having logic models or theories of change that influence what outcomes they are interested in measuring. Many logic models and theories of change also include intermediate outcomes (mediators) or factors that might explain variations in outcomes for different groups of participants (moderators). These mediators or moderators can often help programs understand the context for the longer-term outcomes for which they want to be held accountable. For example, the degree of participants' attendance in and engagement with a program might play a role in whether they develop the skills that are the program's focus. In a workforce program, mediators like certification or skill development might lead to a job.

Many programs that had a logic model or theory of change noted that the model or theory was developed early in the history of the program or at some other key strategic juncture. Staff may have used it initially to determine what outcomes to measure, but it no longer served that purpose. Nor did staff appear to use it to guide program implementation or data collection for ongoing quality improvement. They were more likely to use it in grant applications or to explain the program to funders.

Usability

Program leaders want to prioritize data that will inform how the program is run, how it is performing, which youth are benefitting and which are not, and what might make it better. One program leader put it this way:

"Before you invest in the outcome, know what you're measuring. [...] If data comes back and says, 'no improvements in caring,' what are [program staff] going to do about that?...How are you truly using data for improvement?...If you're not using it to showcase impact and you're not using it for program improvement and it's just sitting in the corner of your office, it's not that helpful."

For some programs, determining what data is usable is a process of trial and error that involves collecting one set of data, discovering that it raises new questions, and then adjusting accordingly. One program described the process as a "journey" in which they have developed a better understanding of how much

money and time it takes to collect particular pieces of data, how to explain the importance of data to staff, and how to tweak what they are measuring so that staff are able to act on the information.

Another organization spoke of the importance of collecting data that is valuable not only to program staff but also participants and avoiding data collection that comes off as critical of youth:

“Evaluation is also instructional, so we try to use student-friendly language so that students who do take the time to digest [the question/s] get some guidance on things that they could be working on. [T]hose are framed developmentally, so that students don’t feel we are judging them. [...] We try to focus, too, on behaviors and attitudes that are malleable, because we want students to feel like they are growing.”

For many programs, data collection can be a source of frustration for staff—either because the data they collect isn’t usable or the logistics of collecting it are burdensome. Staff with [Our Piece of the Pie](#), a multi-service program that provides career readiness services, noted that every contract they manage comes with its own reporting requirements. On their wishlist was a system that would allow them to enter data using their phones while they are already working with youth so they wouldn’t have to schedule a separate time for data collection and entry.

Shareability with participants and community

Many program staff also spoke of the importance of sharing data with youth, parents, schools, and funders. Data sharing—with attention to privacy requirements—can serve a number of purposes, including engaging families and the community and helping stakeholders and program participants track their own progress. The New York City Administration for Children’s Services (ACS NYC), an agency that serves youth involved in the city’s juvenile justice system collects portfolios of students’ work while they are in detention. It shares these portfolios with the students’ home school and the transition specialist. This builds trust between ACS NYC, the transition specialists, and the school system and helps the students establish themselves at their home schools.

Another primary purpose of data sharing is to communicate program impact and effectiveness, which can help with fundraising. Staff at [Mikva Challenge](#) spoke about wanting to tell the stories of their youth participants:

“We really want to show that young people are being transformed by our programming. That they are becoming more civically engaged. Not only when they are part of our program but carrying that on into adulthood. [...] We want to be storytellers because our stories are amazing.”

Program staff felt that stories of this kind were more compelling (and better illustrated young people’s agency in achieving desired outcomes) than the time- and resource-intensive quantitative reports that some funders require.

Youth voice

Incorporating youth voice into selecting outcomes and data collection methods was a priority for many programs in our study. In general, arts and juvenile justice-focused programs did a good job collecting youth feedback by conducting regular check-ins to 1) make sure participants were getting what they needed from the program; 2) build participants’ communication and other SEL skills; 3) regularly gauge participants’ progress; and 4) engage participants in making meaning of data or assessing data collection strategies.

Programs mentioned that soliciting youth voice also means learning about participants’ personal goals and development. For example, a young person in a career readiness program may be interested in a job in a

specific field. Staff in arts, juvenile justice, and workforce programs emphasized that working with youth to identify and take steps toward a goal that is meaningful to them is, in many cases, a more important outcome than the goal itself. In this sense, youth voice is not simply a process to get to an outcome but an outcome itself. Data collection in this instance is about recording participants' goals and progress toward those goals, typically using a database or customer relationship management system. Staff at the Our Piece of the Pie career readiness program explained how the process works:

"[Participants] can have a short-term goal—[which] they have to complete every thirty days—an intermediate goal— [which] might be three to six months—and then a yearly goal. So, we are basically tracking what they are saying they want to achieve and how they are progressing towards that... [W]hen you go and set that goal, you can put action steps within that and set dates you hope they hit, and you can check them off as you go."

Capacity and effort required

Most organizations do not have staff dedicated to data collection and analysis. This responsibility is usually one of many and consequently can end up as a lower priority. One organization mentioned this directly: *"I wrote yesterday to the Board about our need to hire a fulltime Director of Impact... There's so many different things...that we wish we were measuring, and we just don't have the capacity for it."*

In some cases, lack of capacity can force programs to make hard choices about what to measure. As one program put it:

"Our problem is we measure a lot—a lot a lot—and we need to focus a bit more... We're trying to narrow down the five core outcomes we want to see...[We need] staffing, staffing to do the research, staffing to figure it out, staffing to implement it, staffing to analyze it. You know, we sometimes have to say, we don't want to do this survey because we're not going to be able to analyze it for another year, and that's not helpful."

In the absence of dedicated personnel and resources, data collection can be a significant strain on staff, participants, other stakeholders from whom the program collects data, and program budgets. One program called this "survey fatigue," but the issue is much broader than that label suggests. Some programs reported that staff will stop collecting data or even quit their jobs if the process becomes too burdensome or alienates youth. Most staff members who work directly with youth got into the field because they care about them deeply. Being forced to choose between spending their time engaging with young people in fun, productive ways or collecting data (particularly data they do not see being used) can lead to dissatisfaction and burnout. One career readiness program told us:

"I've had staff leave because they just got frustrated with the data, the stuff that was forced on them. And I've had staff say, 'I feel rude when I'm talking to youth and I say, 'I'm going to type this up as we're talking,' because they feel like if they don't get it into the system [they'll] forget to go back or won't get accurate [data]. And they don't like working like that because these people generally are people-persons who want to work with individuals, and they want to give them their attention."

Program participants can also experience burnout because of excessive data collection. This burnout can compromise the quality of the data if participants are not willing to cooperate in the process. An even greater concern is that young people, particularly those who have been marginalized at school or part of the juvenile justice system, may be re-traumatized by intrusive data collection. One program leader said, *"It can be a lot. We've been working with the program staff to balance what is needed versus what is not needed, just so we don't get [youth] burnout. We've really worked to, [in] the last couple of years, scale back where we can or consolidate where we can."*

Program leaders must try to strike the right balance between protecting valuable staff time and program dollars on one hand and ensuring that the program has usable, high-quality data to work with on the other. As with identifying usable data, this balancing act can be a continuous process with programs trying one approach and then adapting it to work better for everyone involved. A staff member at Wings for Kids spoke of the adjustments the program had made in its approach to data collection:

"We tried to do it much more regularly, and it just is overwhelming. You know, we have staff trying to do a Devereux Student Strengths Assessment (DESSA) on 12 kids, and we found that when they were doing it more often, they were just kind of answering and not really putting real thought into the actual assessment of the kids."

Validity and reliability of measures

In order to effectively collect and use data, programs need access to measures that are valid (i.e., accurate) and reliable (i.e., yield consistent results). This can be challenging for two reasons. First, proprietary measures developed by researchers can be prohibitively expensive for many programs. Second, researchers don't always report on the validity or reliability of the measures they develop, and many staff members don't have the training to assess the quality of a given measure on their own.

Mikva Challenge, a civic engagement program in Chicago and Washington D.C., uses measures developed by the [Center for Information and Research on Civic Learning and Engagement \(CIRCLE\)](#) at Tufts University because they have undergone statistical analysis for validity. These measures are not publicly available, but Mikva Challenge has been able to use them because CIRCLE is interested in helping civic engagement organizations determine what to measure and how to measure it and in testing those measures. Organizations that are not well-connected enough to form a partnership of this kind may struggle to find high-quality measures.

Equity

As programs have become more interested in understanding whether they have disparate effects on different groups of participants, they have put more thought into how to identify these disparities and their causes. Many of the programs in the study did this by breaking down the data they collected by race and other demographics. Some civic engagement and arts programs developed outcomes specifically related to equity, such as increased knowledge of how power and inequity operate in society.

Because equity continues to be a topic of rising importance, programs' thinking about it is changing from year to year. It will be important to revisit the question of how OST programs measure equity in 3-4 years as they begin to settle on and codify their approaches.

Chapter 4: Measurement Challenges

What measurement challenges did OST programs experience?

The programs we spoke with noted several measurement challenges. Some were specific to the organization, while others were driven by external factors, such as COVID-19. The measurement challenges programs reported fell into three categories: 1) lack of available or tested measurement tools; 2) difficulty measuring specific types of outcomes; and 3) the burden associated with data collection.

Lack of available or tested measurement tools

Few OST programs reported that they lacked the tools they needed to measure the outcomes of interest to them. The ones that did report a lack of tools identified cultural competence, appreciation for diversity, and changes in anti-racist practices as outcomes they were unable to measure with the tools available. An arts program said, “[we] would love to measure our equity mindset, [and] how are we developing it.” Another program said it would “love to measure [the] impact of celebrating and appreciating youth’s culture in programs.”

Some programs were unaware of research-backed surveys and assessments designed to track the development of SEL skills and related outcomes, such as creativity, identity formation, sense of belonging, self-concept, and confidence. One program said it was interested in “soft skills that can’t be measured in a meaningful way,” while another said of this same category of skills, “not sure where [the measurements] are, [but] we are sure they are out there.” It was not within the scope of our current study to review the tools available for measuring particular outcomes, the tools used by programs, or the quality of those tools. We did learn that many programs used untested tools developed by their own staff.

Specific outcomes that are difficult to measure

Long-term outcomes

Many programs expressed interest in tracking outcomes one or more years after participants left the program. For instance, career and workforce development programs were ultimately interested in understanding if the training, job readiness skills, and work experience participants gained through the program led to better outcomes in terms of future employment, earnings, and benefits. STEM programs mentioned wanting to know if program participation resulted in higher rates of high school graduation and college matriculation, influenced college major and/or career choices, and built long-term trust in scientific principles. For civics programs, long-term outcomes of interest included voting or involvement in a political or advocacy campaign. SEL programs had questions about how developing social and emotional skills affects academic achievement; career development; or relationships with friends, families, or partners. Arts programs were interested in whether their participants were more likely to find lasting happiness, achieve their goals, and participate in civic life.

Programs looking to collect this kind of longitudinal data had trouble keeping track of participants after they left the program, lack of access to administrative records, lack of consent to collect long-term data, and limited staff capacity to do the work. One program described the challenge this way:

“If we’re funding this program, what are the long-term [benefits]? Not just, ‘is this child gaining credits toward graduation this semester for youth in our afterschool program, [...] but what does success look like three months, six months, nine months down the road?’ And we’re not doing that right now because we don’t have a system for getting that information from the school district or those families after the kid has graduated from our program. [...] We need more capacity to maintain that contact longer-term and actually see real outcomes.”

Another said, *“We use [the] National Student Clearinghouse to report on college enrollment [and] persistence.”* This program was interested in knowing if its participants were employed and, if so, *“where are they employed, how much are they making, what the job is. That would be the dream. [We have] explored different avenues such as Equifax work numbers... That is hard to get.”*

Behavior versus knowledge

Many programs reported measuring participants’ knowledge attainment but were unable to determine whether and how knowledge affected behavior. For example, a civics program might track what participants learned about voting but not whether they actually voted. An SEL program might measure the development of participants’ communication skills but have no sense of whether they struggle to put those skills into practice in their daily lives.

School outcomes

While some programs that focused on academic achievement conceived of themselves as extracurricular and therefore were not interested in tracking participants’ progress in school, most wanted access to school data. Many programs did not have data-sharing agreements with school districts and other public agencies. As one put it, *“[We] have a plan. [It’s] just sitting on someone’s desk waiting for signature.”* In the absence of formal agreements, programs had to gather the information from the student themselves. Outcomes like promotion to the next grade and high school graduation were relatively easy to measure in this way because programs could ask all participants about them at the same time. Test scores, grades, and behavior were more of a challenge: Participants did not always want to share the information or did not share it in a timely way, making it difficult for programs to act on it effectively. For example, if a program knew that a student failed an algebra test as soon as it happened, it could provide more tutoring or guidance on study habits before that student failed the entire course. In many cases, however, programs missed that opportunity. Many programs said they needed help engaging and forming trusting partnerships with schools and districts, and developing formal data sharing-agreements to meet their measurement objectives.

Relationships between staff and youth

Many of the programs we spoke to wanted to understand whether and how improvements in a staff member’s skills or knowledge were associated with improved youth outcomes. For example, in one civics program, the focus on democratic classrooms—in which students guided discussions and chose topics to focus on—required understanding on the part of the teachers. Analyzing the relationship between staff and youth learning was a challenge because it required high-quality measurement of both and an understanding of statistical modeling.

Data collection challenges

As noted in Chapter 3, data collection can be a burden on staff, who in many cases are not trained to do the work, and program participants. Some programs in our study were reluctant to ask students to fill out more forms given the heavy testing culture they experience in school. They were also concerned that data collection activities took staff away from delivering program content.

Multiple interviewees noted that their program already collected more data than it could use, saying staff lacked the time for data entry and the expertise to manage the collection process, analyze the data, and create the kind of visual presentations that make it usable and shareable. Some pointed to the burden of fulfilling the reporting requirements of various funders, which often involved recording duplicative information using multiple tools and databases.

We spoke to programs in early 2021 when some were operating online due to COVID-19 and figuring out what data they could collect under the circumstances and how they could collect it. One noted an interest in

“capturing things being said on social media about youth’s experiences with program successes.” Another spoke of trying to *“do discourse analysis of students’ writing and online chat artifacts.”* Barriers to online data collection included participants’ lack of stable internet access and privacy protections and concerns related to tracking information online.

Other data collection challenges noted less frequently include the following:

- Issues of parental consent
- A need to incentivize data collection
- Concerns about social desirability bias (i.e., participants giving the answers they think staff want to hear)
- Lack of funding
- Difficulty gathering information from a representative sample of the participant population
- Poor response rates
- Lack of “kid-friendly” data collection tools
- The ceiling effect (i.e., most participants report positive outcomes at the start of the program making it difficult to measure progress)
- Difficulty tracking outcomes for participants who are in the program for a short period of time

Finally, some programs that do have funding to hire dedicated staff for data collection activities noted that turnover can be a challenge. Data teams can be as small as one person, meaning turnover can bring data collection to a grinding halt.

Measuring equity outcomes can be challenging.

Programs that measure equity outcomes (such as eliminating disparities between demographic groups or increasing participants’ knowledge about how power changes social outcomes) often do so in informal ways, such as staff meetings, staff check-ins, and discussions with participants. These informal methods can make it challenging to quantify progress. Some programs, such as [Generation Citizen](#), are working on updating their logic models to more clearly explain how they will know whether they are achieving their equity goals, but for many this work is new and challenging.

Other programs expressed an interest in accessing data that would help them gauge the effectiveness of their approaches to equity. For instance, [Sherwood Forest](#), a summer camp in which young people participate from 1st grade through high school, was interested in knowing if its outreach and retention efforts have yielded equitable results – *“[Are the] graduating class demographics the same as when they entered? Did [we] lose kids who are worse off? Is the class whiter? This isn’t easy to track.”*

Programs can benefit from tapping outside expertise in data collection and use.

For programs looking to, as one put it, “*change the culture around data...to be less fearful of data and more data-driven,*” bringing in outside experts to help facilitate data collection efforts can make a big difference. One program, [Project Morry](#), discussed being part of a fellowship with [Youth INC.](#), a New York City-based, nonprofit organization that helps programs plan for and carry out data collection and analysis. Youth INC. focuses specifically on SEL outcomes but also helps organizations think broadly about data. Project Morry staff told us:

“...it’s a competitive program and an application process. And what they do is help nonprofits with creating systems for measurement and getting yourself set up to really do this on an on-going basis...So, we’ve been working with them for the last two years and through them we have been trying to really focus our organizational culture on measurement and program evaluation... [SEL outcomes were] something that was always intrinsic to what we were trying to accomplish. [Youth INC.] helped us formalize what we were always trying to do and gave us a way to really measure it. ... We’re still on the journey in terms of...understanding what to do with the data once we get it and how to utilize it to actually make programmatic decisions. But we are moving in that direction. We are surveying our kids. We use Hello Insight, which is an online tool which has made it really easy to do surveys twice a year and get the data and the analytics.”

Chapter 5: Key Findings and Recommendations

The programs in this study made it clear that they cared deeply about measuring participant outcomes and program quality and using what they learned to make improvements and achieve their missions. While many of the staff members we interviewed got into the OST field to work with young people, not data, their desire to help young people develop new skills and achieve success and well-being motivated them to learn a lot about how to collect and use data. This bodes well for the continued commitment of OST programs to measurement and evaluation.

At the same time, the burdens of, and obstacles to, data collection and use that they described were wide-ranging and daunting relative to the resources available to them. Programs need more time, financial support, training, and dedicated staff to ensure that the data they collect is high-quality and that they can act on it effectively.

Summary of findings

The following key findings address the three research questions that were the starting point for the study:

Research question 1: *What is the **range of intended outcomes** for children and youth attending OST programs, including, but not limited to, programs that focus on the arts, civic engagement and social justice, career and workforce development, and general or other services?*

- Programs in the study measured outcomes that were closely related to the content they delivered (e.g., arts, career and workforce readiness, civic engagement and social justice). They also generally measured SEL outcomes and other outcomes required by funders.
- Programs thought critically about how to measure equity outcomes. Approaches to measuring equity included disaggregating data by race and other demographic variables and tracking the development of equity-related knowledge and skills.
- Programs used several criteria to decide what outcomes to track, including consistency with the program's logic model or theory of change, usability and shareability of the data, the effort and capacity required to collect the data, availability of valid and reliable measures, youth interest, and whether measuring a given outcome would promote equity.
- In addition to outcomes, programs consistently measured outputs (i.e., steps that lead to desired outcomes), particularly participation and program quality. This is consistent with findings from our literature scan that public and private funders have invested heavily in quality assessment tools and participation tracking systems. Comparatively few programs used qualitative methods to understand the factors that contribute to program participation and quality.

Research question 2: *What are the **different measurement approaches and instruments** used to document OST program outcomes, including formative and summative measures?*

- Programs reported using a variety of traditional methods to measure outcomes, including *quantitative* methods—like administering surveys and questionnaires, conducting formal assessments, and tracking attendance rates—and *qualitative* methods like conducting interviews and focus groups with young people. Their use of quantitative methods was well-documented, use of qualitative methods less so.
- Programs also used nontraditional methods, such as regular check-ins with participants and reviewing participant journals and portfolios. As with their use of qualitative measures, use of these nontraditional methods was not well-documented.
- Programs found creative, informal ways of tracking SEL outcomes, including games and award systems.

Research question 3: *What are the **gaps in and barriers to documentation and measurement** where intended outcomes are either not measured or are not articulated because of a lack of a measurement strategy, skills, resources, or instruments?*

- The programs we interviewed did not consistently identify specific outcomes of interest that they were unable to measure, although some said they lacked the tools they needed to track equity- or SEL-related outcomes or program quality.
- Programs pointed to broader types of outcomes that were challenging to measure, including longer-term outcomes like college matriculation, career attainment, and participation in civic life; behavior change (e.g., do participants in a civics program exercise their right to vote); school outcomes like test scores and grades that require a data-sharing agreement with the district; and the relationship between improvements in staff members skills and knowledge and youth outcomes.
- Programs reported that the process of collecting data could be burdensome for participants, who have their fill of testing at school, and staff, who in many cases do not have the training or time to do the work. Some pointed specifically to the burden of fulfilling the reporting requirements of various funders, which often involved recording duplicative information using multiple tools and databases.

Recommendations

The following recommendations are divided into three main categories: recommendations for OST programs, recommendations for funders of OST programs, and recommendations for further research. All three sets of recommendations address the challenges experienced by the programs in the study and, in some cases, were suggested by the interviewees themselves.

Recommendations for OST programs

These recommendations for OST programs call for a shift in organizational priorities and approach as opposed to additional funding. In some cases, though, programs may benefit from technical assistance to help them successfully adopt these ideas.

Select a small number of high-priority outcomes to measure.

Programs often cast a wide net for data and end up not using what they collect, which can be frustrating for staff and participants. To avoid this, programs should focus on data they will be able to use and share with participants, families, and the broader community. They may also want to identify outcomes they consider “exploratory”—that is, outcomes they value but choose not to measure for practical or other reasons.

Ensure that the selected outcomes fit into the program’s logic model or theory of change.

Many programs reported that they did not use their logic model or theory of change to decide what outcomes to measure or guide their data collection efforts. Programs should review the outcomes they target for consistency with their logic model or theory of change and, if necessary, make changes to one or the other.

Document the type, source, and purpose of each piece of data collected.

The programs in our study were more likely to keep a record of the data they collected using traditional, quantitative methods like surveys or formal assessments than data collected using qualitative or nontraditional methods like check-ins with participants or portfolios of their work. Programs should thoroughly document all the different ways they collect data to ensure that all data gets used and to better understand what each method contributes to their work.

Make room for youth perspectives in deciding what outcomes to measure and how.

Ultimately, OST programs collect data to better serve young people. Therefore, it is important that programs find out what young people want to gain from their participation (e.g., to learn a particular skill or to get a job in a particular field) and their progress toward those goals. Programs can solicit input from participants not only in deciding what outcomes to measure but also in assessing how well the data collection process is working and in making meaning of what's collected.

Incorporate measures of diversity, equity, and inclusion into data collection efforts.

Programs interested in tracking their progress on issues of equity should consider disaggregating data by race or other demographic variables to identify disparities in participant outcomes; incorporating outcomes specifically related to equity (e.g., the development of civic knowledge and social justice skills) into their logic models or theories of change; revising their curricula to promote those outcomes; and selecting qualitative and, if possible and appropriate, quantitative methods of measuring those outcomes.

Assess the capacity and effort required to collect data and take steps to minimize burnout.

Programs should expect finding the right approach to data collection for staff and participants to be an ongoing process of trial and error. In addition to selecting a small number of high-priority outcomes (as recommended above), programs can help reduce the burden on staff by scheduling data collection for a few specific points in the calendar when data will be most relevant rather than trying to collect it continuously over the course of the program.

Recommendations for funders

To truly get the most out of data collection efforts, OST programs need outside support in a number of forms: additional funding, training and technical assistance, access to high-quality tools, and opportunities to work with researchers and learn from peers. The following recommendations touch on the ways funders can partner with programs on their data journeys.

Connect programs to additional training in key aspects of data collection and use.

Many staff members who oversee their programs' data efforts are trained in youth development and are learning to work with data on the job. Interviewees spoke of the need for more training in assessing which measurement tools are valid and reliable; in developing new databases to replace systems they describe as "clunky" and "cumbersome" or better understand how to navigate the systems they have; and in using data to support program improvement.

Create networks of programs to provide informal technical assistance.

One program in our study secured a grant to create a network of OST programs that could support and learn from one another. Through this network, the program has learned how various types of programs have solved common problems and has worked with other programs' technical assistance specialists to troubleshoot its own particular challenges. This program noted that it would also benefit from participating in networks focused on developing specific skills (e.g., collecting longitudinal data, identifying outcomes of interest and ways to measure them). Conferences serve this function to some extent, but many programs cannot send multiple staff members to conferences due to travel costs or the time it would take away from their regular responsibilities. Virtual meetings or webinars that give participants the opportunity to engage with one another on specific topics of interest are one alternative to in-person conferences.

Provide funding for additional staffing.

Tight budgets mean that a) programs are often unable to hire dedicated staff for data collection and analysis and b) the staff they do have are too busy providing all the services youth in the program need to give data work the attention it demands. Additional resources would allow programs to hire staff whose primary responsibility is data or to reduce staff-to-participant ratios so all relevant staff members have more time for data activities.

Facilitate partnerships with researchers.

Partnerships between practitioners and researchers can be mutually beneficial. Practitioners need guidance on what to measure and how to measure it, while researchers need to field-test the measurement tools they develop or real-world data to analyze. Mivka's Challenge for example has been able to use measurement tools developed by Tufts University, and Tufts researchers have been able to see how those tools work in practice. Research partnerships may also involve formal initiatives in which researchers provide program staff with technical assistance in identifying what outcomes to measure, selecting valid and reliable measurement tools, and analyzing the data they collect.

Make reporting requirements as flexible as possible.

Programs sometimes see their own goals for data collection as being at odds with the data reporting that funders require of them. Almost every program in our study mentioned at least once that meeting the requirements of their various funders was a time- and resource-intensive process that involved recording information that was duplicative or not meaningful to them using multiple tools and databases. Funders should be sensitive to the strain that data collection can place on programs and be as flexible as possible in terms of how and in what format grantees demonstrate progress toward outcomes of interest. As one interviewee put it, *"[If you hired someone to do a job at your house, to build a deck, and then you said, 'Okay, I'll hire you because I like you and the work you do, but I want you to use only my tools,' it would be kind of ridiculous to go about it that way.]"*

Support digital data collection.

As many programs moved online during the COVID-19 pandemic so did their data collection efforts. Even as they return to serving young people in person, programs could benefit from additional funding for, and training in, developing mobile data collection processes (Cava Tadik et al., 2019). Staff at one program noted that the ability to enter data on their phones while they are already working with youth would eliminate the need to schedule a separate time for data collection and entry.

Recommendations for further research

As is often the case with research, this study raised additional questions the answers to which would help inform the field. These recommendations are aimed not only at researchers, but also at funders since the recommendations require resources to carry out.

Expand the existing study to include more programs and voices.

The current study could be expanded by including a) more programs in each of the categories we prioritized; b) programs in additional categories (e.g., STEM, physical fitness, identity development) to increase the generalizability of the findings; and c) perspectives from different types of people involved with OST programs, particularly youth participants.

Further explore how programs make decisions about their data efforts.

Data collection and use require program leaders to make difficult tradeoffs in terms of money and staff time. To help programs navigate these tricky decisions, future research could focus on the following questions:

- How do programs identify outcomes meaningful to them and prioritize those outcomes?
- How do they choose the most appropriate measurement tools for their needs?
- How do they balance spending on improved data collection with spending on programming?
- What types of support for data collection and use (e.g., training, financial, staffing) do they find most helpful?

Highlight innovative solutions to common problems.

This study focused in part on challenges OST programs face in collecting and using data. Future research could help programs address these challenges by identifying “bright spots”—innovative solutions to common problems from exemplary programs. Topics could include forming successful data-sharing partnerships, minimizing the burden of data collection on staff, building staff capacity to collect and analyze data, and securing funding for data efforts.

Investigate strategies for collecting longitudinal data.

Programs in our study expressed interest in tracking long-term participant outcomes but had difficulty doing so. Many programs specifically mentioned the challenge of collecting data from alumni who may no longer be connected to the program. Future research could identify strategies for effectively measuring long-term outcomes by conducting a formal review of best practices in the field and gathering qualitative evidence from programs that have done it successfully.

Catalog existing measurement tools and develop and test new tools as needed.

Through our literature scan, we identified a robust set of tools available for measuring program quality and social and emotional learning, yet some of the programs in our study were unaware of these tools. Researchers could help address this disconnect by cataloging the tools already in use in the field (those developed by researchers and other outside experts and those developed by programs themselves) and assessing their validity and reliability.

The literature scan also revealed a lack of tools relevant to other types of outcomes, including content-specific and equity-related outcomes. Researchers could develop new tools to fill this gap. Funders could support this work by putting out a call for research in field publications like *The Journal for Youth Development*; organizing panels or conference sessions focused on this topic; or forming a network of practitioners, funders, researchers, and content experts to identify and share promising approaches. Funders could also spearhead dissemination efforts to get new and existing tools into the hands of programs that stand to benefit from them. New and existing tools should be tested to ensure they are appropriate for the diverse populations OST programs serve.

Appendices

Appendix I. Additional References

- Augustine, C. H., Sloan McCombs, J., Pane, F.H., Schwartz, H.L., et.. al. (2016). Learning from summer: Effects of voluntary summer learning programs on low-income urban youth. RAND Corporation. <https://www.wallacefoundation.org/knowledge-center/pages/learning-from-summer-effects-of-voluntary-summer-learning-programs-on-low-income-urban-youth.aspx>
- Child Trends. (2020). Programs for youth and young adults: Science-informed definitions. Child Trends. <https://www.childtrends.org/programs-for-youth-and-young-adults-science-informed-definitions>
- Every Hour Counts. (2021). Putting data to work for young people: A framework for measurement, continuous improvement, and equitable systems. <https://wallacefoundation.org/sites/default/files/2023-08/Every-Hour-CountsMeasurement-Framework.pdf>
- Gill, J. (2018). READS: Helping children become summer bookworms. The Wallace Foundation. <https://www.wallacefoundation.org/knowledge-center/pages/reads-helping-children-become-summer-bookworms.aspx>
- Hennink, M., Kaiser, B.N. (2021) Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*. <https://doi.org/10.1016/j.socscimed.2021.114523>.
- Lippman, L., Moore, K.A., Guzman, L., Ryberg, R., McIntosh, H., Ramos, M. F., Caal, S., Carle, A., Kuhfeld, M. (2014). Flourishing children: Defining and testing indicators of positive development. Child Trends. <https://www.childtrends.org/publications/flourishing-children-defining-and-testing-indicators-of-positive-development>
- Lippman, L.H., Moore, K.A. & McIntosh, H. (2011). Positive Indicators of Child Well-Being: A Conceptual Framework, Measures, and Methodological Issues. *Applied Research Quality Life*. 6(1), 425–449. <https://doi.org/10.1007/s11482-011-9138-6>
- Lippman, L.H., Ryberg, R., Terzian, M., Moore, K.A., Humble, J., McIntosh H. (2014). Positive and Protective Factors in Adolescent Well-Being. In: Ben-Arieh A., Casas F., Frønes I., Korbin J. (eds) *Handbook of Child Well-Being*. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-9063-8_141
- Lofland, J., Snow, D., Anderson, L., & Lofland, L.H. (2006). *Analyzing social settings: A guide to qualitative observation and analysis* (4). Wadsworth, Belmont CA.
- McClanahan, W. & Hartmann, T. (2017). *Raising the barre and stretching the canvas: Implementing high quality arts programming in a National Youth Serving Organization*. Philadelphia, PA: Research for Action & McClanahan Associates: <https://www.wallacefoundation.org/knowledge-center/pages/raising-the-barre-and-stretching-the-canvas.aspx>
- McClanahan, W. & Hartmann, T. (2018). *Designing for engagement: The experiences of tweens in the Boys & Girls Clubs' Youth Arts Initiative*. Research for Action. <https://www.wallacefoundation.org/knowledge-center/pages/designing-for-engagement-the-experiences-of-tweens-in-the-boys-and-girls-clubs%E2%80%99-youth-arts-initiative.aspx>
- Montgomery, D., Rogouin, P., & Persuad, N. (2013). *Something to say: Success principles for afterschool arts programs from urban youth and other experts*. The Wallace Foundation. <https://www.wallacefoundation.org/knowledge-center/pages/something-to-say-success-principles-for-afterschool-arts-programs.aspx>

- Neild, R.C., Wilson, S.J., & McClanahan, W. (2019). Afterschool programs: A review of evidence under the Every Student Succeeds Act. Philadelphia: Research for Action.
<https://www.wallacefoundation.org/knowledge-center/Documents/Afterschool-Programs-A-Review-of-Evidence-Under-the-Every-Student-Succeeds-Act.pdf>
- Parise, L. M., Corrin, W., Granito, K., Haider, Z., Somers, M., & Cerna, O. (2017). *Two years of case management for at-risk students: Final findings from the communities in schools random assignment evaluation*. MDRC. <https://www.wallacefoundation.org/knowledge-center/pages/cis-two-years-of-case-management-for-at-risk-students.aspx>
- Redd, Z., Boccanfuso, C., Walker, K., Princiotta, D., Knewstubb, D., & Moore, K. A. (2012). Expanded time for learning both inside and outside the classroom: A review of the evidence base. (Report). Washington, DC: Child Trends. <https://www.childtrends.org/publications/expanded-time-for-learning-both-inside-and-outside-the-classroom-a-review-of-the-evidence-base-executive-summary>
- Russell, C. & Butler, A. (2020). *The afterschool leadership landscape: Supporting and strengthening racial equity*. Washington, DC: Policy Studies Associates.
- Seiden, Katherine (2000). *Development and validation of the "Organizational Readiness for Evaluation" survey instrument*. Doctoral Dissertation, University of Minnesota. Bell & Howell Information and Learning Company. UMI.
- Sloan McCombs, J., Augustine, C. H., Pane, F.H., & Schweig, J. (2020). Every summer counts: A longitudinal analysis of outcomes from the National Summer Learning Project. RAND Corporation. <https://www.wallacefoundation.org/knowledge-center/pages/every-summer-counts-a-longitudinal-analysis-of-outcomes-from-the-national-summer-learning-project.aspx>
- Sloan McCombs, J., Pane, F.H., Augustine, C. H., Schwartz, H. L., et..al. (2014). *Ready for Fall? Near-term effects of voluntary summer learning programs on low-income students' learning opportunities*. RAND Corporation. <https://www.wallacefoundation.org/knowledge-center/pages/ready-for-fall.aspx>
- Somers, M. & Haider, Z. (2017). *Using integrated student supports to keep kids in school: A quasi-experimental evaluation of communities in schools*. MDRC. <https://www.wallacefoundation.org/knowledge-center/pages/cis-using-integrated-student-supports-to-keep-kids-in-school.aspx>
- Sparr, M., Frazier, S., Morrison, C., Miller, K., & Bartko, W.T. (2020). Afterschool programs to improve social-emotional, behavioral, and physical health in middle childhood: A targeted review of the literature. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation & Office on Women's Health, Office of the Assistant Secretary for Health, U.S. Department of Health and Human Services. https://aspe.hhs.gov/sites/default/files/private/pdf/265236/4_MCASP_LiteratureReview.pdf
- Somers, M., Welbeck, R., Grossman, J. B., & Gooden, S. (2015). *An analysis of the effects of an academic summer program for middle school students*. MDRC. <https://www.wallacefoundation.org/knowledge-center/pages/an-analysis-of-the-effects-of-an-academic-summer-program-for-middle-school-students.aspx>
- Wan, Y., J. Ludwig, M., & Boyle, A. (2018). *Review of evidence: Arts education research through the lens of the Every Student Succeeds Act*. American Institute for Research. <https://www.wallacefoundation.org/knowledge-center/pages/review-of-evidence-arts-education-research-essa.aspx>

Appendix II. Bibliography for OST Measurement Literature Scan

Bibliography for OST Measurement Literature Scan

- Abraczinkas M., & Zarret N. (2020). Youth Participatory Action Research for Health Equity: Increasing Youth Empowerment and Decreasing Physical Activity Access Inequities in Under-resourced Programs and Schools. *American Journal of Community Psychology, 66*, 232-243.
- Acar, D., Tertemiz, N., & Taşdemir, A. (2018). The Effects of STEM Training on the Academic Achievement of 4th Graders in Science and Mathematics and their Views on STEM Training Teachers. *International Electronic Journal of Elementary Education, 10*(4), 505–513.
- Ahl, P., Clanton, L., Bitterman, A., Hill, E., Leary, A., Allen, A., & Lawhon, B. (2020). Bigfoot Inspires Youth: Leave No Trace in Urban After-School Programs. *Journal of Outdoor Recreation, Education, and Leadership, 12*(2), 267-.
- Alhassan S., Nwaokelemeh O., Greever C. J., Burkart S., Ahmadi M., St. Laurent C.W., & Barr-Anderson D.J. (2018). Effect of a culturally-tailored mother-daughter physical activity intervention on pre-adolescent African-American girls' physical activity levels. *Preventative Medicine Reports, 11*(7), 7-14.
- Allen, P. J., Chang, R., Gorrall, B. K., Waggenspack, L., Fukuda, E., Little, T. D., & Noam, G. G. (2019). From Quality to Outcomes: A National Study of Afterschool STEM Programming. In *International Journal of STEM Education* (Vol. 6).
- Allen, S., & Peterman, K. (2019). Evaluating Informal STEM Education: Issues and Challenges in Context. *New Directions for Evaluation, 2019*(161), 17–33.
- Allen-Handy, A., Thomas-EL, S.L. & Sung, K.K. (2020). Urban Youth Scholars: Cultivating Critical Global Leadership Development through Youth-Led Justice-Oriented Research. *The Urban Review, 53*(1).
- Anderson, R. E., Jones, S. C. T., Navarro, C. C., McKenny, M. C., Mehta, T. J., & Stevenson, H. C. (2018). Addressing the Mental Health Needs of Black American Youth and Families: A Case Study from the EMBRace Intervention. *International Journal of Environmental Research and Public Health, 15*(5), 898-.
- Annesi, J. J., Smith, A. E., Walsh, S. M., Mareno, N., & Smith, K. R. (2016). Effects of an after-school care-administered physical activity and nutrition protocol on body mass index, fitness levels, and targeted psychological factors in 5- to 8-year-olds. *Translational Behavioral Medicine, 6*(3), 347–357.
- Annesi, J.J., Walsh, S.M., Greenwood, B.L., Mareno, N. and Unruh-Rewkowski, J.L. (2017), Effects of the Youth Fit 4 Life physical activity/nutrition protocol on body mass index, fitness and targeted social cognitive theory variables in 9- to 12-year-olds during after-school care. *Journal of Paediatrics and Child Health, 53*, 365-373
- Annesi, J. J., Walsh, S. M., & Greenwood, B. L. (2016). Increasing Children's Voluntary Physical Activity Outside of School Hours Through Targeting Social Cognitive Theory Variables. *Journal of Primary Care & Community Health, 7*(4), 234–241.
- Annesi J. J. (2019). Generalizability of an adult-validated model for predicting increased physical activity in after-school program children: Effects of treatment formats on fitness indicators. *Evaluation and Program Pplanning, 73*, 33–43.
- Anyon, Y., Kennedy, H., Durbahn, R., & Jenson, J. M. (2018). Youth-Led Participatory Action Research: Promoting Youth Voice and Adult Support in Afterschool Programs. In *Afterschool Matters*.

- Baek, Y., Wang, S., Yang, D., Ching, Y.-H., Swanson, S., & Chittoori, B. (2019). Revisiting Second Graders' Robotics with an Understand/Use-Modify-Create (U²MC) Strategy. In *European Journal of STEM Education* (Vol. 4).
- Baker, S. K., Kamata, A., Wright, A., Farmer, D., & Nippert, R. (2019). Using propensity score matching to estimate treatment effects of afterschool programs on third-grade reading outcomes. *Journal of Community Psychology, 47*(1), 117–134.
- Bang, K. S., Kim, S., Song, M. K., Kang, K. I., & Jeong, Y. (2018). The Effects of a Health Promotion Program Using Urban Forests and Nursing Student Mentors on the Perceived and Psychological Health of Elementary School Children in Vulnerable Populations. *International Journal of Environmental Research and Public Health, 15*(9), 1977.
- Bayless, S. D., Jenson, J. M., Richmond, M. K., Pampel, F. C., Cook, M., & Calhoun, M. (2018). Effects of an Afterschool Early Literacy Intervention on the Reading Skills of Children in Public Housing Communities. In *Child & Youth Care Forum* (Vol. 47).
- Beesley, A. D., Fancsali, C., & Gulemetova, M. (2018). Building Mathematical Identity after School: Year 1 of a Cluster-Randomized Trial. In *Society for Research on Educational Effectiveness*.
- Beets, M. W., Weaver, R. G., Turner-McGrievy, G., Huberty, J., Ward, D. S., Freedman, D., Hutto, B., Moore, J. B., & Beighle, A. (2016). Making Healthy Eating Policy Practice: A Group Randomized Controlled Trial on Changes in Snack Quality, Costs, and Consumption in After-School Programs. *American Journal of Health Promotion: AJHP, 30*(7), 521–531.
- Beets, M. W., Turner-McGrievy, B., Weaver, R. G., Huberty, J., Moore, J. B., Ward, D. S., & Freedman, D. A. (2015). Intervention leads to improvements in the nutrient profile of snacks served in afterschool programs: A group randomized controlled trial. *Translational Behavioral Medicine, 3*(3), 329–338.
- Beets, M. W., Weaver, R. G., Turner-McGrievy, G., Huberty, J., Moore, J. B., Ward, D. S., Freedman, D. A., & Beighle, A. (2017). Two-Year Healthy Eating Outcomes: An RCT in Afterschool Programs. *American Journal of Preventive Medicine, 53*(3), 316–326.
- Bertrand, M. (2018). Youth Participatory Action Research and Possibilities for Students of Color in Educational Leadership. In *Educational Administration Quarterly* (Vol. 54).
- Beymer, P. N., Rosenberg, J. M., Schmidt, J. A., & Naftzger, N. J. (2018). Examining Relationships among Choice, Affect, and Engagement in Summer STEM Programs. *Journal of Youth and Adolescence, 47*(6), 1178–1191.
- Blagoeva, N. V., Karppinen, S., & Kairavuori, S. (2019). The Integrated Approach to Teaching Visual Art in After-School Activity Classes. In *International Journal of Art & Design Education* (Vol. 38).
- Blazer, C. (2016). After-School Academic Enrichment Programs. Information Capsule. Volume 1509. In *Research Services, Miami-Dade County Public Schools*.
- Branscum, P., Housley, A., Bhochhibhoya, A., & Hayes, L. (2016). A Formative Evaluation of Healthy Heroes: A Photo Comic Book-Social Cognitive Theory Based Obesity Prevention Program. In *Journal of Health Education Teaching* (Vol. 7).
- Bremer, E., Graham, J. D., Bedard, C., Rodriguez, C., Kriellaars, D., & Cairney, J. (2020). The Association between Playfun and Physical Activity: A Convergent Validation Study. In *Research Quarterly for Exercise and Sport* (Vol. 91).
- Brock, L. L., Murrah, W. M., Cottone, E. A., Mashburn, A. J., & Grissmer, D. W. (2017). An After-School Intervention Targeting Executive Function and Visuospatial Skills Also Improves Classroom Behavior. In *International Journal of Behavioral Development* (Vol. 42).
- Broder, E. D., Guilbert, K. E., Tinghitella, R. M., Murphy, S. M., Ghalambor, C. K., & Angeloni, L. M. (2019). Authentic Science with Dissemination Increases Self-Efficacy of Middle School Students. *Integrative and Comparative Biology, 59*(6), 1497–1508.

- Brown, A. A., Outley, C. W., & Pinckney, H. P. (2018). Examining the Use of Leisure for the Sociopolitical Development of Black Youth in Out-of-School Time Programs. *Leisure Sciences, 40*(7), 686–696.
- Browne, L., & Warner, R. (2019). Being Present in the Moment. *The Camping Magazine, 92*(5), 30–35.
- Budd, E. L., Nixon, C. T., Hymel, A. M., & Tanner-Smith, E. E. (2020). The impact of afterschool program attendance on academic outcomes of middle school students. *Journal of Community Psychology, 48*(8), 2439–2456.
- Burack, C., Melchoir, A., & Hoover, M. (2019). Do After-School Robotics Programs Expand the Pipeline into STEM Majors in College? In *Journal of Pre-College Engineering Education Research* (Vol. 9).
- Bussert-Webb K., & Henry L. A. (2017). Promising Digital Practices for Nondominant Learners. *International Journal of Educational Technology, 4*(2), 43–55.
- Camasso, M. J., & Jagannathan, R. (2018). Nurture thru Nature: Creating Natural Science Identities in Populations of Disadvantaged Children through Community Education Partnership. In *Journal of Environmental Education* (Vol. 49).
- Cappella, E., Hwang, S. H. J., Kieffer, M. J., & Yates, M. (2018). Classroom Practices and Academic Outcomes in Urban Afterschool Programs: Alleviating Social-Behavioral Risk. In *Journal of Emotional and Behavioral Disorders* (Vol. 26).
- Cavendish, L. M., Vess, S. F., & Li-Barber, K. (2016). Collaborating in the Community: Fostering Identity and Creative Expression in an Afterschool Program. In *Journal of Language and Literacy Education* (Vol. 12).
- Chaddock-Heyman, L., Erickson, K. I., Kienzler, C., Drollette, E. S., Raine, L. B., Kao, S. C., Bensken, J., Weissshappel, R., Castelli, D. M., Hillman, C. H., & Kramer, A. F. (2018). Physical Activity Increases White Matter Microstructure in Children. *Frontiers in Neuroscience, 12*, 950.
- Chaddock-Heyman, L., Weng, T. B., Kienzler, C., Weissshappel, R., Drollette, E. S., Raine, L. B., Westfall, D. R., Kao, S. C., Baniqued, P., Castelli, D. M., Hillman, C. H., & Kramer, A. F. (2020). Brain Network Modularity Predicts Improvements in Cognitive and Scholastic Performance in Children Involved in a Physical Activity Intervention. *Frontiers in Human Neuroscience, 14*, 346.
- Chambers, Salter, & Muldrow. (2019). Getting Past the Gateway: An Exploratory Case on Using Utilitarian Scientific Literacy to Support First-Year Students At-Risk of Leaving STEM. *Education Sciences, 9*(4), 265-.
- Champine, R. B., Wang, J., Ferris, K. A., Hershberg, R. M., Erickson, K., Johnson, B. R., & Lerner, R. M. (2016). Exploring the Out-of-School Time Program Ecology of Boy Scouts. *Research in Human Development, 13*(2), 97–110.
- Chan, H., Choi, H., Hailu, M. F., Whitford, M., & Duplechain DeRouen, S. (2020). Participation in structured STEM-focused out-of-school time programs in secondary school: Linkage to postsecondary STEM aspiration and major. *Journal of Research in Science Teaching, 57*(8), 1250–1280.
- Ching, Y.-H., Yang, D., Wang, S., Baek, Y., Swanson, S., & Chittoori, B. (2019). Elementary School Student Development of STEM Attitudes and Perceived Learning in a STEM Integrated Robotics Curriculum. In *TechTrends: Linking Research and Practice to Improve Learning* (Vol. 63).
- Chittum, J. R., Jones, B. D., Akalin, S., & Schram, Á. B. (2017). The Effects of an Afterschool STEM Program on Students' Motivation and Engagement. In *International Journal of STEM Education* (Vol. 4).
- Clevenger, K. A., Moore, R. W., Sutton, D., Montoye, A. H. K., Trost, S. G., & Pfeiffer, K. A. (2018). Accelerometer Responsiveness to Change between Structured and Unstructured Physical Activity in Children and Adolescents. In *Measurement in Physical Education and Exercise Science* (Vol. 22).

- Coelho, V. A., & Sousa, V. (2018). Differential Effectiveness of a Middle School Social and Emotional Learning Program: Does Setting Matter?. *Journal of Youth and Adolescence*, 47(9), 1978–1991.
- Coimbra, S., Catarino, C., Nascimento, H., Inês Alves, A., Filipa Medeiros, A., Bronze-da-Rocha, E., Costa, E., Rocha-Pereira, P., Aires, L., Seabra, A., Mota, J., Ferreira Mansilha, H., Rêgo, C., Santos-Silva, A., & Belo, L. (2017). Physical exercise intervention at school improved hepcidin, inflammation, and iron metabolism in overweight and obese children and adolescents. *Pediatric Research*, 82(5), 781–788.
- Coles, J. A. (2019). The Black Literacies of Urban High School Youth Countering Antiblackness in the Context of Neoliberal Multiculturalism. In *Journal of Language and Literacy Education* (Vol. 15).
- Colman, A. (2018). Parks and Recreation: A Champion for YOUTH Health. *Parks & Recreation (Arlington, Va.)*, 53(6), 54–59.
- Colt, A., & Reilly, J. M. (2019). An Educational Intervention to Improve the Sleep Behavior and Well-Being of High School Students. *PRIMER (Leawood, Kan.)*, 3, 21.
- Cown, M. H., Grossman, B. M., & Giraudo, S. Q. (2017). Nutrition Education Intervention to Improve Nutrition-Related Knowledge, Attitudes, and Behaviors for Hispanic Children. *Ecology of Food and Nutrition*, 56(6), 493–513.
- Cradock AL, Barrett JL, Giles CM, Lee RM, Kenney EL, deBlois ME, Thayer JC, Gortmaker SL (2016). Promoting Physical Activity With the Out of School Nutrition and Physical Activity (OSNAP) Initiative: A Cluster-Randomized Controlled Trial. *JAMA Pediatr.* 2016 Feb;170(2):155-62.
- Cromer, Kelly & D'Agostino, Emily & Hansen, Eric & Alfonso, Caitlin & Frazier, Stacy. (2019). After-school poly-strengths programming for urban teens at high-risk for violence exposure. *Translational Behavioral Medicine*, 9(3).
- Crouter, S. E., Salas, C., & Wiecha, J. (2017). Effects of an afterschool community center physical activity program on fitness and body composition in obese youth. *Journal of Sports Sciences*, 35(11), 1034–1040.
- Cryan, M., & Martinek, T. (2017). Youth Sport Development through Soccer: An Evaluation of an After-School Program Using the TPSR Model. In *Physical Educator* (Vol. 74).
- Cutucache, C., Boham, T., Luhr, J., Sommers, A., Stevenson, N., Sointu, E., Mäkitalo-Siegl, K., Kärkkäinen, S., Valtonen, T., Grandgenett, N., & Tapprich, W. (2018). NE STEM 4U Afterschool Intervention Leads to Gains in STEM Content Knowledge for Middle School Youth. In *Cogent Education* (Vol. 5).
- D'Agostino, E. M., Frazier, S. L., Hansen, E., Nardi, M. I., & Messiah, S. E. (2020). Association of a Park-Based Violence Prevention and Mental Health Promotion After-School Program With Youth Arrest Rates. *JAMA Network Open*, 3(1), e1919996.
- D'Agostino, E. M., Frazier, S. L., Hansen, E., Patel, H. H., Ahmed, Z., Okeke, D., Nardi, M. I., & Messiah, S. E. (2019). Two-Year Changes in Neighborhood Juvenile Arrests After Implementation of a Park-Based Afterschool Mental Health Promotion Program in Miami-Dade County, Florida, 2015–2017. *American Journal of Public Health*, 109(53), S214–S220.
- D'Agostino, E. M., Patel, H. H., Ahmed, Z., Hansen, E., Sunil Mathew, M., Nardi, M. I., & Messiah, S. E. (2018). Impact of change in neighborhood racial/ethnic segregation on cardiovascular health in minority youth attending a park-based afterschool program. *Social Science & Medicine* (1982), 205, 116–129.
- D'Agostino, E. M., Patel, H. H., Hansen, E., Mathew, M. S., Nardi, M., & Messiah, S. E. (2018). Longitudinal analysis of cardiovascular disease risk profile in neighbourhood poverty subgroups: 5-year results from an afterschool fitness programme in the USA. *Journal of Epidemiology and Community Health*, 72(3), 193–201.

- D'Agostino, E. M., Patel, H. H., Hansen, E., Mathew, M. S., Nardi, M. I., & Messiah, S. E. (2018). Effect of participation in a park-based afterschool program on cardiovascular disease risk among severely obese youth. *Public Health, 159*, 137–143.
- Daniel, S. M., & Eley, C. (2018). Improving Cohesion in Our Writing: Findings from an Identity Text Workshop with Resettled Refugee Teens. In *Journal of Adolescent & Adult Literacy* (Vol. 61).
- Danifo, J., & Valdez, J. (2019). Building Social Emotional Learning in Youth Through Humanities-Centered Activities. *Young Adult Library Services, 17*(2), 27–31.
- Dauenhauer, B., Keating, X., & Lambdin, D. (2016). Effects of a Three-Tiered Intervention Model on Physical Activity and Fitness Levels of Elementary School Children. *The Journal of Primary Prevention, 37*(4), 313–327.
- Davis, C. L., Litwin S. E., Pollock N. K., Waller J. L., Haidong Z., Dong Y., Kapuku G., Bhagatwala J., Harris R. A., Looney J., Williams C. F., Armento A., Schmidt M. D., & Bassali R. (2020). Exercise effects on arterial stiffness and heart health in children with excess weight: The SMART RCT. *International Journal of Obesity, 44*, 1152–1163. Davis, J. N., Martinez, L. C., Spruijt-Metz, D., & Gatto, N. M. (2016). LA Sprouts: A 12-Week Gardening, Nutrition, and Cooking Randomized Control Trial Improves Determinants of Dietary Behaviors. *Journal of Nutrition Education and Behavior, 48*(1), 2–11.e1.
- Delacruz, S., & Guerra, P. (2019). Building Sustainable Afterschool Literacy Programs by Partnering with University Teacher Candidates. In *School Community Journal* (Vol. 29).
- Demetriou Y., Reimers A. K., Alesi M., Scifo L., Chicau Borrego C., Monteiro D., & Kelso A. (2019). Effects of school-based interventions on motivation towards physical activity in children and adolescents: protocol for a systematic review. *Systematic Reviews, 8*, 113.
- Duberg, A., Jutengren, G., Hagberg, L., & Möller, M. (2020). The effects of a dance intervention on somatic symptoms and emotional distress in adolescent girls: A randomized controlled trial. *The Journal of International Medical Research, 48*(2), 300060520902610.
- Dunkerly-Bean, J. M., Bean, T. W., Kidd, D., & Johnson, E. (2017). “What Color Are Our Hearts?” Challenging Social and Literacy Inequalities in an Elementary School Writing Club. In *AERA Online Paper Repository*.
- Durlak, J. A. (2017). The Fundamental Importance of Effective Program Implementation for Successful Character Development. *Journal of Character Education, 13*(2), 1–11.
- Eckert, S., Metzger-Riftkin, J., & Nurmis, J. (2018). Teaching Girls Online Skills: Results of the Wikid Grrls Intervention. In *Journal of Media Literacy Education* (Vol. 10).
- Economos, C. D., Hennessy, E., Chui, K., Dwyer, J., Marcotte, L., Must, A., Naumova, E. N., & Goldberg, J. (2020). Beat osteoporosis - nourish and exercise skeletons (BONES): a group randomized controlled trial in children. *BMC Pediatrics, 20*(1), 83.
- Eisenhauer, S. (2018). The Micro Temporal Arc: A Practical Planning Tool for Afterschool Student Engagement. In *Afterschool Matters*.
- Engell, T., Kirkøen, B., Hammerstrøm, K. T., Kornør, H., Ludvigsen, K. H., & Hagen, K. A. (2020). Common Elements of Practice, Process and Implementation in Out-of-School-Time Academic Interventions for At-risk Children: A Systematic Review. *Prevention Science, 21*(4), 545–556.
- Erwin, H. E., Rose, S. A., Small, S. R., & Perman, J. (2016). Jumpin’ Jaguars: Encouraging Physical Activity After School. In *Afterschool Matters*.
- Farrell, A. F., Collier-Meek, M. A., & Furman, M. J. (2019). Supporting Out-of-School Time Staff in Low Resource Communities: A Professional Development Approach. *American Journal of Community Psychology, 63*(3–4), 378–390.

- Faust, L., & Kuperminc, G. P. (2020). Psychological Needs Fulfillment and Engagement in Afterschool: "I Pay Attention Because I Am Really Enjoying This." In *Journal of Adolescent Research* (Vol. 35).
- Fettig, A., Cook, A. L., Morizio, L., Gould, K., & Brodsky, L. (2018). Using dialogic reading strategies to promote social-emotional skills for young students: An exploratory case study in an after-school program. *Journal of Early Childhood Research, 16*(4), 436–448.
- Frahm, R. (2016). The Promise: A Case Study of Say Yes to Education in Buffalo. Principles for Effective Education Grantmaking. Case Study No. 15. In *Grantmakers for Education*.
- Franklin, C. (2017). Effects of the Afterschool Program on Student Achievement of Students with Disabilities in a Rural Georgia Middle School. In *ProQuest LLC*.
- Frazier, S. L., Rusch, D., Coxe, S., Stout, T. J., Helseth, S. A., Dirks, M. A., Bustamante, E. E., Atkins, M. S., Glisson, C., Green, P. D., Bhaumik, D., & Bhaumik, R. (2020). After-School Programs and Children's Mental Health: Organizational Social Context, Program Quality, and Children's Social Behavior. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, 1–14. Advance online publication.
- Frosini, C. (2017). An "I" in Teen? Perceived Agency in a Youth Development Program. In *Afterschool Matters*.
- Gagnon, R. J. (2020). Examining perceived adolescent socioemotional development and repeated camp experiences using a planned missing data design. *Journal of Leisure Research, 51*(5), 517–535.
- Gagnon, R. J., Garst, B. A., & Townsend, J. A. (2019). Tough decisions in medical specialty camps: Relationships between camp dosage, outcomes, and camper attendance. *Social Science & Medicine (1982), 221*, 49–57.
- Gagnon, R. J., & Sandoval, A. (2020). Pre-college STEM camps as developmental context: Mediational relations between gender, career decidedness, socioemotional development, and engagement. *Children and Youth Services Review, 108*, 104584-.
- Garcia, I., Grossman, J. B., Herrera, C., Strassberger, M., Dixon, M., & Linden, L. (2020). Aiming Higher: Assessing Higher Achievement's Out-of-School Expansion Efforts. In *MDRC*.
- Gatz, J. (2018). Middle School Girls' Science Motivation and Performance: Cognitive Effects of an Out-of-School Time Program with Nutrition and Fitness Components. *ProQuest LLC*.
- Gatz, J., & Kelly, A. M. (2018). Afterschool School Triathlon Training for 11- to 14-Year Old Girls: Influences on Academic Motivation and Achievement. In *Health Education Journal* (Vol. 77).
- Gatz, J., Kelly, A. M., & Clark, S. L. (2019). Improved Executive Function and Science Achievement for At-Risk Middle School Girls in an Aerobic Fitness Program. *The Journal of Early Adolescence, 39*(3), 453–469.
- Germeroth, C., Kelleman, B., & Spartz, J. (2018). Lyrics2Learn: Teaching Fluency through Music and Technology. *Education Sciences, 8*(3), 91-.
- Gibson, J. J. (2017). An Assessment of Factors Relating to High School Students' Science Self-Efficacy. In *ProQuest LLC*.
- Gilliam, M., Bouris, A., Hill, B., & Jagoda, P. (2016). "The Source": An Alternate Reality Game to Spark STEM Interest and Learning among Underrepresented Youth. *Journal of STEM Education, 17*(2), 14-.
- Gilliam, M., Jagoda, P., Fabiyi, C., Lyman, P., Wilson, C., Hill, B., & Bouris, A. (2017). Alternate Reality Games as an Informal Learning Tool for Generating STEM Engagement among Underrepresented Youth: A Qualitative Evaluation of the Source. *Journal of Science Education and Technology, 26*(3), 295–308.

- Giraud, S. Q., Rivera-Gonzalez, N., Berg, A., Hartzell, D., & Grossman, B. M. (2019). Nutrition Intervention to Hispanic Groups: Pilot Studies with Children and Caregivers. *Ecology of Food and Nutrition, 58*(3), 219–235.
- Goff, J. E. (2016). Building Stronger Communities: The Reciprocity between University, Student, and Community through Service-Learning. In *ProQuest LLC*.
- Golan, M. & Wiessam, A. A. (2018) School-based versus after-school delivery of a universal wellness programme – A randomized controlled multi-arm trial. *Eating Behaviors, 31*, 41–47.
- Goldenberg, B. M. (2016). Youth Historians in Harlem: An After-School Blueprint for History Engagement through the Historical Process. In *Social Studies* (Vol. 107).
- Gomoll, A., Hmelo-Silver, C. E., Šabanovic, S., & Francisco, M. (2016). Dragons, Ladybugs, and Softballs: Girls' STEM Engagement with Human-Centered Robotics. In *Journal of Science Education and Technology* (Vol. 25).
- Gorard, S., Siddiqui, N., See, B. H., Smith, E., & White, P. (2017). Children's University: Evaluation Report and Executive Summary. In *Education Endowment Foundation*.
- Gordon, B., Jacobs, J. M., & Wright, P. M. (2016). Social and Emotional Learning through a Teaching Personal and Social Responsibility Based After-School Program for Disengaged Middle-School Boys. In *Journal of Teaching in Physical Education* (Vol. 35).
- Grack Nelson, A., Goeke, M., Auster, R., Peterman, K., & Lussenhop, A. (2019). Shared Measures for Evaluating Common Outcomes of Informal STEM Education Experiences. *New Directions for Evaluation, 2019*(161), 59–86.
- Grack Nelson, A. L. (2017). Development and Validation of a Survey to Measure Perceived Team Communication Skills in Middle and High School STEM Out-of-School Time Programs. *ProQuest LLC*.
- Graves, J. M., Mackelprang, J. L., Barbosa-Leiker, C., Miller, M. E., & Li, A. Y. (2017). Quality of life among working and non-working adolescents. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 26*(1), 107–120.
- Gray, C. (2017). Investigating African-American Parent Perception of Successful After-School Programs in Urban School Districts: A Case Study. In *ProQuest LLC*.
- Gray, P., Rule, A. C., Kirkland Holmes, G., Logan, S. R., Alert, A. L., & Mason, C. A. (2016). Learning Form and Function by Dance-Dramatizing Cultural Legends to Drum Rhythms Wearing Student-Made Animal Masks. In *Journal of STEM Arts, Crafts, and Constructions* (Vol. 1).
- Green Jr., E. M. (2018). Students' Perceptions of the Effect of an Afterschool Program on Their Path to College. In *ProQuest LLC*.
- Greenberg Motamedi, J., & Singh, M. (2016). Oregon MESA: Improving Grades in Science and Mathematics. In *Education Northwest*.
- Gross, M., Latham, D., Underhill, J., & Bak, H. (2016). The Peritext Book Club: Reading to Foster Critical Thinking about STEAM Texts. In *School Library Research* (Vol. 19).
- Gutierrez de Blume, A. P. (2017). The Effects of Strategy Training and an Extrinsic Incentive on Fourth- and Fifth-Grade Students' Performance, Confidence, and Calibration Accuracy. In *Cogent Education* (Vol. 4).
- Habibi, A., Damasio, A., Ilari, B., Elliott Sachs, M., & Damasio, H. (2018). Music training and child development: a review of recent findings from a longitudinal study. *Annals of the New York Academy of Sciences, 10.1111/nyas.13606*. Advance online publication.
- Haegele, J. A., & Poretta, D. L. (2017). A Theory-Based Physical Education Intervention for Adolescents with Visual Impairments. In *Journal of Visual Impairment & Blindness* (Vol. 111).

- Harper, S. G. (2017). Engaging Karen Refugee Students in Science Learning through a Cross-Cultural Learning Community. In *International Journal of Science Education* (Vol. 39).
- Hartmann, T., & McClanahan, W. (2020). Designing for Engagement: How High-Quality Arts OST Programs Can Engage Tweens. In *Afterschool Matters*.
- Hathaway, D. A. (2018). Afterschool Program Effectiveness on Literacy Growth for Children Attending a High Poverty School. In *ProQuest LLC*.
- Hauseman, D. C. (2016). Youth-led community arts hubs: Self-determined learning in an out-of-school time (OST) program. *Cogent Education*, 3(1).
- Hava, K., Guyer, T., & Cakir, H. (2020). Gifted Students' Learning Experiences in Systematic Game Development Process in After-School Activities. In *Educational Technology Research and Development* (Vol. 68).
- Hayes, S. (2018). Engaging Girls in STEM. *AMLE Magazine*, 6(2), 26–28.
- Hedemann, E. R., & Frazier, S. L. (2017). Leveraging After-School Programs to Minimize Risks for Internalizing Symptoms Among Urban Youth: Weaving Together Music Education and Social Development. *Administration and Policy in Mental Health*, 44(5), 756–770.
- Heerman, W. J., Elsakary, Y., Sommer, E. C., Escarfuller, J., & Barkin, S. L. (2020). Assessing the scale and spread of an experiential teaching kitchen in after-school programming among school-age children. *Public Health Nutrition*, 1–8. Advance online publication.
- Helmick, M., Esmond, A. C., Hedrick, V., Zoellner, J., You, W., & Hill, J. L. (2019). The Adoption of the Healthy Eating Standards in Local Afterschool Programs Does Not Improve Quality of Snacks. In *Journal of School Health* (Vol. 89).
- Helseth, S. A., & Frazier, S. L. (2017). Peer-Assisted Social Learning for Diverse and Low-Income Youth: Infusing Mental Health Promotion Into Urban After-School Programs. *Administration and Policy in Mental Health and Mental Health Services Research*, 45(2), 286–301.
- Hennessy, S. L., Sachs, M. E., Ilari, B., & Habibi, A. (2019). Effects of Music Training on Inhibitory Control and Associated Neural Networks in School-Aged Children: A Longitudinal Study. *Frontiers in neuroscience*, 13, 1080.
- Heo, M., Jimenez, C. C., Lim, J., Isasi, C. R., Blank, A. E., Lounsbury, D. W., Fredericks, L., Bouchard, M., Faith, M. S., & Wylie-Rosett, J. (2018). Effective nationwide school-based participatory extramural program on adolescent body mass index, health knowledge and behaviors. *BMC pediatrics*, 18(1), 7.
- Hill, E., Williams, R., McIntosh, T., Morris, D., Hill, L., & Duncan, M. (2020). The Impact of Out of School Time Triathlon Camps on Perceived Competence, Interests in Exploration, and Responsibility among Youth Campers: A Pilot Study. *Recreation, Parks, and Tourism in Public Health*, 4, 33–41.
- Hill, J. C., Lynne-Landsman, S. D., Graber, J. A., & Johnson, K. J. (2016). Evaluating a Pregnancy and STI Prevention Programme in Rural, At-Risk, Middle School Girls in the USA. In *Health Education Journal* (Vol. 75).
- Holmes, V. R. (2016). Stacey and Bo Porter SELF Foundation Afterschool Program: Academic and Social Impacts on Key, Revere, and Meyerland Middle Schools, 2015-2016. Evaluation Report. Volume 11, Issue 1. In *Houston Independent School District*.
- Holmes, V. R. (2018). District III – East Area Fine Arts Initiative: Year 2 – Building Capacity for Quality Fine Arts in Schools, 2017-2018. Research Educational Program Report. In *Houston Independent School District*.
- Hopkins, M., Provenzano, A. M., & Spencer, M. S. (2016). Benefits, Challenges, Characteristics and Instructional Approaches in an El Sistema Inspired After-School String Program Developed as a University-School Partnership in the United States. In *International Journal of Music Education* (Vol. 35).

- Houck, C. D., Barker, D. H., Hadley, W., Brown, L. K., Lansing, A., Almy, B., & Hancock, E. (2016). The 1-year impact of an emotion regulation intervention on early adolescent health risk behaviors. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association, 35*(9), 1036–1045.
- Houck, C. D., Barker, D. H., Hadley, W., Menefee, M., & Brown, L. K. (2018). Sexual Risk Outcomes of an Emotion Regulation Intervention for At-Risk Early Adolescents. *Pediatrics, 141*(6), e20172525.
- Houck, C. D., Hadley, W., Barker, D., Brown, L. K., Hancock, E., & Almy, B. (2016). An Emotion Regulation Intervention to Reduce Risk Behaviors Among at-Risk Early Adolescents. *Prevention Science: The Official Journal of the Society for Prevention Research, 17*(1), 71–82.
- Hsu, P.-S., Lee, E. M., Ginting, S., Smith, T. J., & Kraft, C. (2019). A Case Study Exploring Non-Dominant Youths' Attitudes toward Science through Making and Scientific Argumentation. In *International Journal of Science and Mathematics Education* (Vol. 17).
- Huang, D., Leon, S., & La Torre, D. (2017). Using Entropy Balancing to Reduce the Effects of Selection Bias in Afterschool Studies: An Example in Studying the Relationship between Intensity of Afterschool Program Participation and Academic Achievement. In *International Journal for Research on Extended Education* (Vol. 5).
- Huang, H., Fernandez, S. B., Rhoden, M.-A., & Joseph, R. (2019). Elements of a College Support Program That Matter: A Case Study. *Research on Social Work Practice, 29*(8), 949–960.
- Ige, T. J., DeLeon, P., & Nabors, L. (2017). Motivational Interviewing in an Obesity Prevention Program for Children. *Health Promotion Practice, 18*(2), 263–274.
- Isoldi, K. K., & Dolar, V. (2016). Cooking Up Energy with Predominately Latino Children during Afterschool Hours. *American journal of health behavior, 40*(5), 634–644.
- Johnson, J. A. (2019). The Effect of Online Cross-Age Peer Tutoring on Student Self-Efficacy in Middle School STEM. In *ProQuest LLC*.
- Jones, J. S., Milton, F., Mostazir, M., & Adlam, A. R. (2020). The Academic Outcomes of Working Memory and Metacognitive Strategy Training in Children: A Double-Blind Randomized Controlled Trial. In *Developmental Science* (Vol. 23).
- Kahan, D., & McKenzie, T. L. (2018.). Physical Activity and Energy Expenditure during an After-School Running Club: Laps versus Game Play. In *Journal of School Health* (Vol. 88).
- Kahan, D., & McKenzie, T. L. (2018). Physical Activity and Energy Expenditure During an After-School Running Club: Laps Versus Game Play. *The Journal of School Health, 88*(3), 237–245.
- Kaufman, C. E., Schwinn, T. M., Black, K., Keane, E. M., Big Crow, C. K., Shangreau, C., Tuitt, N. R., Arthur-Asmah, R., & Morse, B. (2018.). Impacting Precursors to Sexual Behavior among Young American Indian Adolescents of the Northern Plains: A Cluster Randomized Controlled Trial. In *Journal of Early Adolescence* (Vol. 38).
- Ke, F., Clark, K. M., & Uysal, S. (2019). Architecture Game-Based Mathematical Learning by Making. In *International Journal of Science and Mathematics Education* (Vol. 17).
- Kelley, M. N., & Lowe, J. R. (2018). A Culture-Based Talking Circle Intervention for Native American Youth at Risk for Obesity. *Journal of Community Health Nursing, 35*(3), 102–117.
- Kennedy, H., Matyasic, S., Schofield Clark, L., Engle, C., Anyon, Y., Weber, M., Jimenez, C., Osiemo Mwirigi, M., & Nisle, S. (2020). Early Adolescent Critical Consciousness Development in the Age of Trump. In *Journal of Adolescent Research* (Vol. 35).
- Khalil, G. E., Wang, H., Calabro, K. S., Mitra, N., Shegog, R., & Prokhorov, A. V. (2017). From the Experience of Interactivity and Entertainment to Lower Intention to Smoke: A Randomized Controlled Trial and Path Analysis of a Web-Based Smoking Prevention Program for Adolescents. *Journal of Medical Internet Research, 19*(2), e44.

- Khalil, G. E., Wang, H., Calabro, K. S., & Prokhorov, A. V. (2019). Revealing users' experience and social interaction outcomes following a web-based smoking prevention intervention for adolescents: A qualitative study. *PLoS one*, *14*(10), e0223836.
- Kinder, C. J., Gaudreault, K. L., Jenkins, J. M., Wade, C. E., & Woods, A. M. (2019). At-Risk Youth in an After-School Program: Structured vs. Unstructured Physical Activity. In *Physical Educator* (Vol. 76).
- King, T. C., & Garcia, M. (2016). Examining the Relationship between After-School Persistence and an Early Warning Index Associated with Dropout. In *AERA Online Paper Repository*.
- Kirkpatrick, B. A., Wright, S., Daniels, S., Taylor, K. L., McCurdy, M., & Skinner, C. H. (2-19). Tootling in an After-School Setting: Decreasing Antisocial Interactions in At-Risk Students. In *Journal of Positive Behavior Interventions* (Vol. 21).
- Kostina-Ritchey, E., Velez-Gomez, P., & Dodd, S. L. (2017). Student Assets and Commitment to Learning in an Afterschool Leadership Development Program: Looking beyond the Myths. In *Middle Grades Research Journal* (Vol. 11).
- Koydemir, S., Şahin, İ., Durhan, S., Oğan, Ş., Gözükar, C., & Çokluk, G. (2017). Changing High School Students' Attitudes Towards Mathematics in a Summer Camp: Happiness Matters. *Educational Sciences: Theory & Practice*, *17*(5), 1625–1648.
- Kulkin, M. (2016). Math Is Like a Scary Movie? Helping Young People Overcome Math Anxiety. In *Afterschool Matters*.
- La Torre, D., Leon, S., Wang, J., & Cai, L. (2019). Long-Term Outcome Study: A Longitudinal Study of LA's BEST Students' Persistence and Graduation Rates. CRESST Report 862. In *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.
- Lagesse, R., & Marshall, D. (2019). Girls in Engineering: How to Increase the Number of Young Girls in Technology and Engineering Courses! In *Technology and Engineering Teacher* (Vol. 78).
- Landau, Z., Abiri, S., Lebenthal, Y., Jakubowicz, D., Mor, N., Lerner-Geva, L., Boaz, M., Wainstein, J., & Bar-Dayana, Y. (2018). Lifestyle intervention program benefits children with overweight compared to children with obesity. *Obesity Research & Clinical Practice*, *12*(1), 85–92.
- Landry, M. J., Markowitz, A. K., Asigbee, F. M., Gatto, N. M., Spruijt-Metz, D., & Davis, J. N. (2019). Cooking and Gardening Behaviors and Improvements in Dietary Intake in Hispanic/Latino Youth. *Childhood Obesity (Print)*, *15*(4), 262–270.
- Langberg, J. M., Evans, S. W., Schultz, B. K., Becker, S. P., Altaye, M., & Girio-Herrera, E. (2016). Trajectories and Predictors of Response to the Challenging Horizons Program for Adolescents With ADHD. *Behavior Therapy*, *47*(3), 339–354.
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2016). Out-of-School-Time Programs: A Meta-Analysis of Effects for At-Risk Students. *Review of Educational Research*, *76*(2), 275–313.
- Leas, H. D., Nelson, K. L., Grandgenett, N., Tappich, W. E., & Cutucache, C. E. (2017). Fostering curiosity, inquiry, and scientific thinking in elementary school students: Impact of the NE STEM 4U intervention. *Journal of Youth Development (Online)*, *12*(2), 103–120.
- Lee, J., Zhang, T., Chu, T., Gu, X., & Zhu, P. (2020). Effects of a Fundamental Motor Skill-Based Afterschool Program on Children's Physical and Cognitive Health Outcomes. *International Journal of Environmental Research and Public Health*, *17*(3), 733.
- Lee, O., Park, M., Jang, K., & Park, Y. (2017). Life lessons after classes: investigating the influence of an afterschool sport program on adolescents' life skills development. *International Journal of Qualitative Studies On Health And Well-Being*, *12*(1), 1307060.

- Lee, D. S., Dang, T. G., Ulibas-Pascual, J., Gordon Biddle, K. A., Heller de Leon, B., Elliott, D., & Gorter, J. (2017). Exploring the Influence of Efficacy Beliefs and Homework Help in Predicting Reading Achievement Among Underserved Children in an Afterschool Program. *The Urban Review*, 49(5), 707–728.
- Lee, R. M., Barrett, J. L., Daly, J. G., Mozaffarian, R. S., Giles, C. M., Cradock, A. L., & Gortmaker, S. L. (2019). Assessing the effectiveness of training models for national scale-up of an evidence-based nutrition and physical activity intervention: A group randomized trial. *BMC Public Health*, 19(1), 1587–1587.
- Liang, Y., Lau, P., Jiang, Y., & Maddison, R. (2020). Getting Active with Active Video Games: A Quasi-Experimental Study. *International Journal of Environmental Research And Public Health*, 17(21), 7984.
- Linder, D. E., Mueller, M. K., Gibbs, D. M., Alper, J. A., & Freeman, L. M. (2018). Effects of an Animal-Assisted Intervention on Reading Skills and Attitudes in Second Grade Students. In *Early Childhood Education Journal* (Vol. 46).
- Lindo, E. J., Weiser, B., Cheatham, J. P., & Allor, J. H. (2018). Benefits of Structured After-School Literacy Tutoring by University Students for Struggling Elementary Readers. In *Grantee Submission* (Vol. 34).
- Little, C. A., Adelson, J. L., Kearney, K. L., Cash, K., & O'Brien, R. (2018). Early Opportunities to Strengthen Academic Readiness: Effects of Summer Learning on Mathematics Achievement. *The Gifted Child Quarterly*, 62(1), 83–95.
- Liu, Y., Simpkins, S. D., & Lin, A. R. (2018). Ethnic Cultural Features in Organized Activities: Relations to Latino Adolescents' Activity Experiences and Parental Involvement. *Journal of Youth and Adolescence*, 47(10), 2243–2260.
- Logan N. E., Raine L. B., Drollette E. S., Castelli D. M., Khan, N. A., Kramer A. F., & Hillman C. H. (2020). The differential relationship of an afterschool physical activity intervention on brain function and cognition in children with obesity and their normal weight peers. *Pediatric Obesity*, 16(2).
- Lu, K. D., Cooper, D. M., Haddad, F., & Radom-Aizik, S. (2018). Four Months of a School-Based Exercise Program Improved Aerobic Fitness and Clinical Outcomes in a Low-SES Population of Normal Weight and Overweight/Obese Children With Asthma. *Frontiers in Pediatrics*, 6, 380.
- Lucko, J. (2018). “We Didn’t Have Courage”: Internalizing Racism and the Limits of Participatory Action Research. In *Anthropology & Education Quarterly* (Vol. 49).
- Luesse, H. B., Luesse, J. E., Lawson, J., Koch, P. A., & Contento, I. R. (2019). “In Defense of Food” Curriculum: A Mixed Methods Outcome Evaluation in Afterschool. In *Health Education & Behavior* (Vol. 46).
- Lynch, A. D., Ferris, K. A., Burkhard, B., Wang, J., Hershberg, R. M., & Lerner, R. M. (2016). Character Development within Youth Development Programs: Exploring Multiple Dimensions of Activity Involvement. *American Journal of Community Psychology*, 57(1–2), 73–86.
- Mac Iver, M. A., & Mac Iver, D. J. (2019). “STEMming” the Swell of Absenteeism in the Middle Years: Impacts of an Urban District Summer Robotics Program. *Urban Education (Beverly Hills, Calif.)*, 54(1), 65–88.
- Mac Iver, M. A., Sheldon, S., Naeger, S., & Clark, E. (2017). Mentoring Students Back On-Track to Graduation: Program Results from Five Communities. *Education and Urban Society*, 49(7), 643–675.
- Marshall, B. P., Loya, A., Drazan, J., Prato, A., Conley, N., Thomopoulos, S., & Reuther, K. (2020). Developing A STEM+M Identity in Underrepresented Minority Youth Through Biomechanics and Sports-Based Education. *Journal of Biomechanical Engineering*, 10.1115/1.4047548. Advance online publication.
- Martin, J. J., Byrd, B., Garn, A., McCaughtry, N., Kulik, N., & Centeio, E. (2016). Predicting Social Responsibility and Belonging in Urban After-School Physical Activity Programs with Underserved Children. In *Urban Review: Issues and Ideas in Public Education* (Vol. 48).

- Maynard, B. R., Solis, M. R., Miller, V. L., & Brendel, K. E. (2017). Mindfulness-based interventions for improving cognition, academic achievement, behavior, and socioemotional functioning of primary and secondary school students. *Campbell Systematic Review*, 13(1), 1-144.
- McClanahan, W. S., & Hartmann, T. A. (2018). Designing for Engagement: The Experiences of Tweens in the Boys & Girls Clubs' Youth Arts Initiative. In *Wallace Foundation*.
- McDaniel, S. C., & Besnoy, K. D. (2019). Cross-Age Peer Mentoring for Elementary Students with Behavioral and Academic Risk Factors. In *Preventing School Failure* (Vol. 63).
- McDaniel, S., & Yarbrough, A.-M. (2016). A Literature Review of Afterschool Mentoring Programs for Children at Risk. In *Journal of At-Risk Issues* (Vol. 19).
- McDougall, M. A., Schaeffer, R., Holm, R., & Specker, B. (2016). Increased Activity in Unstructured Versus Structured Gym Time in an After-School Program. *South Dakota Medicine: The Journal of the South Dakota State Medical Association*, 69(12), 541-549.
- McGovern, J., Drewson, S. R., Hope, A., & Konopack, J. F. (2020). Gender Differences in a Youth Physical Activity Intervention: Movement Levels and Children's Perceptions. In *American Journal of Health Education* (Vol. 51).
- McGrath, R. E. (2018). What Is Character Education? Development of a Prototype. *Journal of Character Education*, 14(2), 23-35.
- McLaughlin, M. W. (2018). You Can't Be What You Can't See: The Power of Opportunity to Change Young Lives. In *Harvard Education Press*.
- Messiah, S. E., D'Agostino, E. M., Hansen, E., Mathew, M. S., Okeke, D., Nardi, M., Kardys, J., & Arheart, K. L. (2018). Longitudinal Impact of a Park-Based Afterschool Healthy Weight Program on Modifiable Cardiovascular Disease Risk Factors in Youth. *Journal of Community Health*, 43(1), 103-116.
- Messiah, S. E., D'Agostino, E. M., Patel, H. H., Hansen, E., Mathew, M. S., & Arheart, K. L. (2019). Changes in cardiovascular health and physical fitness in ethnic youth with intellectual disabilities participating in a park-based afterschool programme for two years. *Journal of Applied Research in Intellectual Disabilities*, 32(6), 1478-1489.
- Messiah, S. E., D'Agostino, E. M., Patel, H. H., Hansen, E., Mathew, M. S., & Arheart, K. L. (2018). Sex differences in fitness outcomes among minority youth after participation in a park-based after-school program. *Annals of Epidemiology*, 28(7), 432-439.
- Messiah, S. E., Jiang, S., Kardys, J., Hansen, E., Nardi, M., & Forster, L. (2016). Reducing childhood obesity through coordinated care: Development of a park prescription program. *World Journal of Clinical Pediatrics*, 5(3), 234-243.
- Messiah, S. E., Vidot, D., Hansen, E., Kardys, J., Sunil Matthew, M., Nardi, M., & Arheart, K. L. (2017). Impact of a park-based afterschool program replicated over five years on modifiable cardiovascular disease risk factors. *Preventive Medicine*, 95, 66-73.
- Montoya, M. A. (2020). Bringing Afterschool Home: OST Programming in Affordable Housing Communities. In *Afterschool Matters*.
- Morano, M., Robazza, C., Bortoli, L., Rutigliano, I., Ruiz, M. C., & Campanozzi, A. (2020). Physical Activity and Physical Competence in Overweight and Obese Children: An Intervention Study. *International Journal of Environmental Research And Public Health*, 17(17), 6370.
- Moreno, N. P., Tharp, B. Z., Vogt, G., Newell, A. D., & Burnett, C. A. (2016). Preparing Students for Middle School through After-School STEM Activities. In *Journal of Science Education and Technology* (Vol. 25).

- Mouza, C., Marzocchi, A., Pan, Y.-C., & Pollock, L. (2016). Development, Implementation, and Outcomes of an Equitable Computer Science After-School Program: Findings from Middle-School Students. In *Journal of Research on Technology in Education* (Vol. 48).
- Mouza, C., Pan, Y.-C., Yang, H., & Pollock, L. (2020). A Multiyear Investigation of Student Computational Thinking Concepts, Practices, and Perspectives in an After-School Computing Program. In *Journal of Educational Computing Research* (Vol. 58).
- Moyer, L., Klopfer, M., & Ernst, J. V. (2018). Bridging the Arts and Computer Science: Engaging At-Risk Students through the Integration of Music. In *Technology and Engineering Teacher* (Vol. 77).
- Muzaffar, H., Nikolaus, C. J., Ogolsky, B. G., Lane, A., Liguori, C., & Nickols-Richardson, S. M. (2019). Promoting Cooking, Nutrition, and Physical Activity in Afterschool Settings. *American Journal of Health Behavior, 43*(6), 1050–1063.
- Nabors, L., Murphy, J. M., Lusky, C., Young, C.-J., & Sanger, K. (2020). Using Photovoice to Improve Healthy Eating for Children Participating in an Obesity Prevention Program. *Global Pediatric Health*.
- Nascimento, H., Alves, A. I., Medeiros, A. F., Coimbra, S., Catarino, C., Bronze-da-Rocha, E., Costa, E., Rocha-Pereira, P., Silva, G., Aires, L., Seabra, A., Mota, J., Ferreira Mansilha, H., Rêgo, C., Santos-Silva, A., & Belo, L. (2016). Impact of a School-Based Intervention Protocol - ACORDA Project - On Adipokines in An Overweight and Obese Pediatric Population. *Pediatric Exercise Science, 28*(3), 407–416.
- Nation, J. M., Harlow, D., Arya, D. J., & Longtin, M. (2019). Being and Becoming Scientists: Design-Based STEM Programming for Girls. In *Afterschool Matters*.
- Neild, R. C., Wilson, S. J., & McClanahan, W. (2019a). A Companion to Afterschool Programs: A Review of Evidence Under the Every Student Succeeds Act. Afterschool Evidence Guide. In *Research for Action*.
- Neild, R. C., Wilson, S. J., & McClanahan, W. (2019b). Afterschool Programs: A Review of Evidence under the Every Student Succeeds Act. In *Research for Action*.
- Ngo, B. (2017). Naming Their World in a Culturally Responsive Space: Experiences of Hmong Adolescents in an After-School Theatre Program. In *Journal of Adolescent Research* (Vol. 32).
- Nguyen, B., & Murimi, M. W. (2017). An After-School Cultural and Age-Sensitive Nutrition Education Intervention for Elementary Schoolchildren. *Journal of Nutrition Education and Behavior, 49*(10), 877–880.e1.
- Nieder, L. E. (2019). Effects of an Academic Enrichment Program on Elementary-Aged Students' Performance. In *ProQuest LLC*.
- Noel, C. R., & Getch, Y. Q. (2016). Noncontingent Reinforcement in After-School Settings to Decrease Classroom Disruptive Behavior for Students with Autism Spectrum Disorder. *Behavior Analysis in Practice, 9*(3), 261–265.
- O'Donnell, J., & Kirkner, S. L. (2016). Helping Low-Income Urban Youth Make the Transition to Early Adulthood: A Retrospective Study of the YMCA Youth Institute. In *Afterschool Matters*.
- Owens, J. C. (2018). Attempting to Reduce Anger: An International Evaluation of a Forgiveness Based Therapeutic Intervention for Students Who Bully Others. In *ProQuest LLC*.
- Paiement, K., Drapeau, V., Gilbert, J. A., Lemoyne, J., Moreau, N., Monthuy-Blanc, J., Tremblay, J., Marcil, V., & Mathieu, M. E. (2020). Changes in Lifestyle Habits among Adolescent Girls after FitSpirit Participation. *International Journal of Environmental Research and Public Health, 17*(12), 4388.
- Paluta, L. M., Lower, L., Anderson-Butcher, D., Gibson, A., & Iachini, A. L. (2020). Examining the Quality of 21st Century Community Learning Center After-School Programs: Current Practices and Their Relationship to Outcomes. In *Children & Schools* (Vol. 38).

- Pan, Y.-C. (2018). Exploring the Impact of a School-University Partnership Model on Supporting Computer Science Learning among Middle School Students. In *ProQuest LLC*.
- Parker, C., Grigg, J., D'Souza, S., Mitchell, C., & Smith, E. (2020). Informed aspirations in science and engineering with upper elementary students after 1 year of a STEM intensive university-school district partnership. *School Science and Mathematics, 120*(6), 364–374.
- Parker, C., Kruchten, C., & Moshfeghian, A. (2017). Connecting Urban Students with Engineering Design: Community-Focused, Student-Driven Projects. In *Afterschool Matters*.
- Parker, M. A., Segovia, E., & Tap, B. (2016). Examining Literature on Hispanic Student Achievement in the Southeastern United States and North Carolina. In *Journal of Hispanic Higher Education* (Vol. 15).
- Pasiali, V., & Clark, C. (2018). Evaluation of a Music Therapy Social Skills Development Program for Youth with Limited Resources. *Journal of Music Therapy, 55*(3), 280–308.
- Patchen, A. K., Patchen, A. K., Zhang, L., Zhang, L., Barnett, M., & Barnett, M. (2017). Growing Plants and Scientists: Fostering Positive Attitudes toward Science among All Participants in an Afterschool Hydroponics Program. *Journal of Science Education and Technology, 26*(3), 279–294.
- Pearce, K., & Dollman, J. (2019). Healthy for Life Pilot Study: A Multicomponent School and Home Based Physical Activity Intervention for Disadvantaged Children. *International Journal of Environmental Research and Public Health, 16*(16), 2935.
- Pentz, M. A., Hieftje, K. D., Pendergrass, T. M., Brito, S. A., Liu, M., Arora, T., Tindle, H. A., Krishnan-Sarin, S., & Fiellin, L. E. (2019). A videogame intervention for tobacco product use prevention in adolescents. *Addictive Behaviors, 91*, 188–192.
- Pinnell, M., Kurup, R., Stock, R., Turney, V., & Wendell, T. (2019). Using Engineering Design to Increase Literacy among Third Graders: A Pilot Study. In *Research in the Schools* (Vol. 25).
- Post, P., & Palacios, R. (2019). Aggie Play: A Gender-Relevant Physical Activity Program for Girls. *Journal of sport & exercise psychology, 41*(4), 194–205.
- Price, C. A., Kares, F., Segovia, G., & Loyd, A. B. (2018). Staff matter: Gender differences in science, technology, engineering or math (STEM) career interest development in adolescent youth. *Applied Developmental Science, 23*(3), 239–254.
- Ramer, J. D., Santiago-Rodríguez, M. E., Davis, C. L., Marquez, D. X., Frazier, S. L., & Bustamante, E. E. (2020). Exercise and Academic Performance Among Children With Attention-Deficit Hyperactivity Disorder and Disruptive Behavior Disorders: A Randomized Controlled Trial. *Pediatric Exercise Science*, 1–10. Advance online publication.
- Riedinger, K., & Taylor, A. (2016). "I Could See Myself as a Scientist": The Potential of Out-of-School Time Programs to Influence Girls' Identities in Science. In *Afterschool Matters*.
- Riiser, K., Haugen, A. L. H., Lund, S., & Løndal, K. (2019). Physical Activity in Young Schoolchildren in After School Programs. In *Journal of School Health* (Vol. 89).
- Robbins, L. B., Ling, J., Sharma, D. B., Dalimonte-Merckling, D. M., Voskuil, V. R., Resnicow, K., Kaciroti, N., & Pfeiffer, K. A. (2019). Intervention Effects of "Girls on the Move" on Increasing Physical Activity: A Group Randomized Trial. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine, 53*(5), 493–500.
- Robbins, L. B., Ling, J., & Wen, F. (2020). Attending After-School Physical Activity Club 2 Days a Week Attenuated an Increase in Percentage Body Fat and a Decrease in Fitness Among Adolescent Girls at Risk for Obesity. *American Journal of Health Promotion: AJHP, 34*(5), 500–504.
- Roberts, G. J., Capin, P., Roberts, G., Miciak, J., Quinn, J. M., & Vaughn, S. (2018). Examining the Effects of Afterschool Reading Interventions for Upper Elementary Struggling Readers. In *Remedial and Special Education* (Vol. 39).

- Roddick, A. (2019). Calls for Coaches: Coaching Social and Emotional Skills in Youth Sports. *The Aspen Institute: Project Play*.
- Rojas, N. (2016). The Association of Peer Behavioral Regulation with School Readiness Skills in Preschool. In *Society for Research on Educational Effectiveness*.
- Roy, B. K. (2017). Cultural Identity and Third Space: An Exploration of Their Connection in a Title I School. In *ProQuest LLC*.
- Rudnick, S. G. (2018). The Effects of a College Awareness Program on Elementary Students' Attitudes toward College. In *ProQuest LLC*.
- Russell, C. A., & Francis, Y. (2018). Evaluation of the YMI Cornerstone Mentoring Program: Role in Supporting Engagement in School and Learning. In *Policy Studies Associates, Inc.*
- Russell, C. A., & Meredith, J. (2019). Literacy Learning in Drop-In Library Programs: Evaluation of the Free Library of Philadelphia's LEAP Program. In *Policy Studies Associates, Inc.*
- Russell, C. A., Turner, T. T., & Butler, A. (2018). Voices of DC Parents and Youth on OST. In *Policy Studies Associates, Inc.*
- Sanchez, J. E., & Usinger, J. (2019). An Evaluation of a Pilot Robotics Program. *Journal of STEM Education*, 20(1), 11–16.
- Sanderson, C. N. (2017). Implementing Visual Culture Curriculum for Underserved Populations in an Afterschool Ceramics Program. In *Online Submission*.
- Schmidt, J. A., Beymer, P. N., Rosenberg, J. M., Naftzger, N. N., & Shumow, L. (2020). Experiences, activities, and personal characteristics as predictors of engagement in STEM-focused summer programs. *Journal of Research in Science Teaching*, 57(8), 1281–1309.
- Schneider, E. (2020). Academic Afterschool Support for English Language Learners in Middle School through a Literacy Methods Course: A Summary of Five Years of Experiences. In *SRATE Journal*(Vol. 29).
- Schnittka, J., & Schnittka, C. (2016). "Can I Drop It This Time?" Gender and Collaborative Group Dynamics in an Engineering Design-Based Afterschool Program. In *Journal of Pre-College Engineering Education Research*(Vol. 6).
- Schultz, B. K., Evans, S. W., Langberg, J. M., & Schoemann, A. M. (2017). Outcomes for adolescents who comply with long-term psychosocial treatment for ADHD. *Journal of Consulting and Clinical Psychology*, 85(3), 250–261.
- Shanahan, L. E., McVee, M. B., Slivestri, K. N., & Haq, K. (2016). Disciplinary Literacies in an Engineering Club: Exploring Productive Communication and the Engineering Design Process. In *Literacy Research: Theory, Method, and Practice*(Vol. 65).
- Shapiro, V. B., Shapiro, V. B., Accomazzo, S., Accomazzo, S., Robitaille, J. L., & Robitaille, J. L. (2017). In the Same Ballpark or a Whole New Ball Game? Staff as Raters of Youth Behavior. *Journal of Child and Family Studies*, 26(4), 1051–1055.
- Sheltzer, J. M., & Consoli, A. J. (2019). Understanding the impact of an after-school music program with engaged underserved youth. *Journal of Community Psychology*, 47(6), 1364–1379.
- Smith, C. D., & Hope, E. C. (2020). "We Just Want to Break the Stereotype": Tensions in Black Boys' Critical Social Analysis of Their Suburban School Experiences. In *Journal of Educational Psychology*(Vol. 112).
- Smith, L. M. (2017). "Book Club Rules, and Tutoring Drools": An Intervention Mixed Methods Study of the Effects of an After-School Book Club on Third-Grade Boys' Reading Achievement, Attitudes, and Preferences. In *ProQuest LLC*.

- Solar, E. (2018). Mindfulness Meditation with High School Students Who Receive Special Education Services: Research Recommendations from a Pilot Study. In *Journal of the American Academy of Special Education Professionals*.
- St. John, A. L. (2018). The Impact of Expanded Access and Increased Engagement and the Effect on Closing the Achievement Gap. In *ProQuest LLC*.
- Staecker, E., Puett, E., Afrassiab, S., Ketcherside, M., Azim, S., Rhodes, D., & Wang, A. (2016). Effectiveness of an Afterschool-Based Aggression Management Program for Elementary Students. In *Professional School Counseling* (Vol. 19).
- Stevens, B., Farrelly, J., & Quell, A. (2017). Integrating Positive Behavioral Interventions and Supports Into an Afterschool Tennis Program for At-Risk Youth. *Report on Emotional & behavioral Disorders in Youth, 17*(2), 47–51.
- Sucre, S. (2016). Teaching the Whole Child through Physical Education and Youth Development. *Strategies (Reston, Va.)*, 29(2), 42-.
- Swanson Hoyle, K. J. (2017). Investigating the Interactions, Beliefs, and Practices of Teacher-Coach Teams in a STEM After-School Setting. In *ProQuest LLC*.
- Thompson, B. J., Stock, M. S., Mota, J. A., Drusch, A. S., DeFranco, R. N., Cook, T. R., & Hamm, M. A. (2017). Adaptations Associated with an After-School Strength and Conditioning Program in Middle-School-Aged Boys: A Quasi-Experimental Design. *Journal of Strength and Conditioning Research, 31*(10), 2840–2851.
- Tichavakunda, A. A. (2019). Fostering College Readiness: An Ethnography of a Latina/o Afterschool Program. In *Education and Urban Society* (Vol. 51).
- Tobin, E. M., & Colley, S. (2018). Getting Back on Track at Twilight. In *Phi Delta Kappan* (Vol. 100).
- Trajković, N., Pajek, M., Sporiš, G., Petrinović, L., & Bogataj, Š. (2020). Reducing Aggression and Improving Physical Fitness in Adolescents Through an After-School Volleyball Program. *Frontiers in Psychology, 11*, 2081.
- Trott, C. D. (2020). Children's Constructive Climate Change Engagement: Empowering Awareness, Agency, and Action. In *Environmental Education Research* (Vol. 26).
- Tymes, D. D., Outlaw, K. L., & Hamilton, B. K. (2016). Life Skills Interventions to Improve Social Confidence, Self-Management, and Protection against Drug Use in Rural Elementary School Aged Children. *Journal of Community Health Nursing, 33*(1), 11–19.
- Updegraff, K. A., Umaña-Taylor, A. J., Rodríguez De Jesús, S. A., McHale, S. M., Feinberg, M. F., & Kuo, S. I. (2016). Family-focused prevention with Latinos: What about sisters and brothers?. *Journal of Family Psychology: JFP: Journal of the Division of Family Psychology of the American Psychological Association (Division 43), 30*(5), 633–640.
- Ustach, E. (2020). Flexibility and Fidelity in a Drop-In, Open-Door Art Studio Program. In *Afterschool Matters*.
- Vandell, D. L., Lee, K., Whitaker, A. A., & Pierce, K. M. (2020). Cumulative and Differential Effects of Early Child Care and Middle Childhood Out-of-School Time on Adolescent Functioning. *Child Development, 91*(1), 129–144.
- Vilkin, E., Einhorn, L., Satyanarayana, S., Eisu, A., Kimport, K., & Flentje, A. (2020). Elementary Students' Gender Beliefs and Attitudes Following a 12-Week Arts Curriculum Focused on Gender. In *Journal of LGBT Youth* (Vol. 17).
- Visker, J., Cox, C., Banez, J. C., Azim, S., Wang, A., Dunsieith, N., & Lasser, B. (2020). Assessing Changes in Attitudes Concerning Developmental Assets among Selected Middle School Students. In *Health Educator* (Vol. 52).

- Walther, A., Chai, W., Dunker, T., Franzen-Castle, L., & Krehbiel, M. (2018). A New Way to Assess Nutrition Knowledge: The Healthy Plate Photo Method. In *Afterschool Matters*.
- Walther, A., Dunker, T., Franzen-Castle, L., & Krehbiel, M. (2018). Increasing At-Risk Youth Self-Reported and Objectively Measured Physical Activity in an Afterschool Program. In *Journal of Family and Consumer Sciences* (Vol. 110).
- Wang, M. L., Otis M., Milagros R. C., Griecci C. F., & Lemon S. C. (2019) Reducing sugary drink intake through youth empowerment: results from a pilot-site randomized study. *International Journal of Behavioral Nutrition and Physical Activity*, 16, 58.
- Wang, J., Champine, R. B., Ferris, K. A., Hershberg, R. M., Warren, D. J., Burkhard, B. M., Su, S., & Lerner, R. M. (2017). Is the Scoutreach Initiative of Boy Scouts of America Linked to Character Development among Socioeconomically, Racially, and Ethnically Diverse Youth?: Initial Explorations. *Journal of Youth and Adolescence*, 46(10), 2230–2240.
- Wang, Y.-C. (2016). Exploring Culturally and Linguistically Diverse Students' Identities in an Afterschool Book Club. In *ProQuest LLC*.
- Wanless, E., Judge, L. W., Dieringer, S. T., & Johnson, J. E. (2017). Recruiting Youth for After-School Health Intervention Programming: Parent and Student Perceptions. In *Physical Educator* (Vol. 74).
- Weaver, R Glenn, Brazendale, K., Chandler, J. L., Turner-McGrievy, G. M., Moore, J. B., Huberty, J. L., Ward, D. S., & Beets, M. W. (2017). First year physical activity findings from turn up the HEAT (Healthy Eating and Activity Time) in summer day camps. *PloS One*, 12(3), e0173791–e0173791.
- Weiss, M. R., Kipp, L. E., Phillips Reichter, A., Espinoza, S. M., & Bolter, N. D. (2019). Girls on the Run: Impact of a Physical Activity Youth Development Program on Psychosocial and Behavioral Outcomes. *Pediatric Exercise Science*, 31(3), 330–340.
- Weiss, M. R., Bolter, N. D., & Kipp, L. E. (2016). Evaluation of The First Tee in promoting positive youth development: Group comparisons and longitudinal trends. *Research Quarterly for Exercise and Sport*, 87(3), 271–283.
- Whitson M. L, Robinson S., Van Valkenburg K., & Jackson M. (2019). The benefits of an afterschool music program for low-income urban youth: The music haven evaluation project. *Journal of Community Psychology*, 48(2), 426–436.
- Whooten, R. C., Horan, C., Cordes, J., Dartley, A. N., Aguirre, A., & Taveras, E. M. (2020). Evaluating the Implementation of a Before-School Physical Activity Program: A Mixed-Methods Approach in Massachusetts, 2018. *Preventing Chronic Disease*, 17, E116–E116.
- Wieland, M. L., Biggs, B. K., Brockman, T. A., Johnson, A., Meiers, S. J., Sim, L. A., Tolleson, E., Hanza, M. M., Weis, J. A., Rosenman, J. R., Novotny, P. J., Patten, C. A., Clark, M. M., Millerbernd, J., & Sia, I. G. (2020). Club Fit: Development of a Physical Activity and Healthy Eating Intervention at a Boys & Girls Club After School Program. *The Journal of Primary Prevention*, 41(2), 153–170.
- Williams, C. F., Bustamante, E. E., Waller, J. L., & Davis, C. L. (2019). Exercise effects on quality of life, mood, and self-worth in overweight children: the SMART randomized controlled trial. *Translational Behavioral Medicine*, 9(3), 451–459.
- Wilson, C. (2016). Impact of an Afterschool Program on Middle School MAP Scale Scores for Math and Communication Arts. In *ProQuest LLC*.
- Wolfe, W. S., & Dollahite, J. (2021). Evaluation of the Choose Health: Food, Fun, and Fitness 3rd- to 6th-Grade Curriculum: Changes in Obesity-Related Behaviors. *The Journal of School Health*, 91(1), 9–18.
- Wolfe, W. S., Scott-Pierce, M., & Dollahite, J. (2018). Choose Health: Food, Fun, and Fitness Youth Curriculum Promotes Positive Behaviors. *Journal of Nutrition Education and Behavior*, 50(9), 924–930.

- Wong, W. W., Ortiz, C. L., Stuff, J. E., Mikhail, C., Lathan, D., Moore, L. A., Alejandro, M. E., Butte, N. F., & Smith, E. O. (2016). A Community-based Healthy Living Promotion Program Improved Self-esteem Among Minority Children. *Journal of Pediatric Gastroenterology and Nutrition*, 63(1), 106–112.
- Worker, S. M., Iaccopucci, A. M., Bird, M., & Horowitz, M. (2018). Promoting Positive Youth Development Through Teenagers-as-Teachers Programs. *Journal of Adolescent Research*, 34(1), 30–54.
- Zimmerman, M. A., Eisman, A. B., Reischl, T. M., Morrel-Samuels, S., Stoddard, S., Miller, A. L., Hutchison, P., Franzen, S., & Rupp, L. (2018). Youth Empowerment Solutions: Evaluation of an After-School Program to Engage Middle School Students in Community Change. In *Health Education & Behavior* (Vol. 45).
- Zyromski, B., Martin, I., & Mariani, M. (2019). Evaluation of the True Goals School Counseling Curriculum: A Pilot Study. In *Journal for Specialists in Group Work* (Vol. 44).

Appendix III. Expert Informants

List of experts that recommended programs to be included in interviews and survey

Expert Name	Organizational Affiliation (at the time of the interview)
Gigi Antoni	The Wallace Foundation
Deana Around Him	Child Trends
Charles Barrios	Administration for Children's Services, NYC
Bronwyn Bevan	The Wallace Foundation
Deb Bialeschki	American Camp Association
Dale Blyth	University of Minnesota
Isaac Castillo	Youth Invest Partners (formerly Venture Philanthropy Partners)
Krista Collins	The David P. Weikart Center for Youth Program Quality, The Forum for Youth Investment
David Condliffe	Center for Community Alternatives
Elizabeth Devaney	The Children's Institute
Sean Flanagan	America's Promise Alliance
Felipe Franco	The Annie E. Casey Foundation
Tracey Hartmann	Research for Action
Heather Ikemire	National Guild for Community Arts Education
Mimi Ito	University of California, Irvine
Ranita Jain	The Aspen Institute
Stephanie Jones	Harvard University
Suzanne Le Menestrel	National Academy of Sciences
Akiva Liberman	Child Trends
Ivan Lui	Brooklyn Bridge Alliance for Youth
Kelly Murphy	Child Trends
Rachel Oberg-Hauser	City of Minneapolis
Jen Rinehart	Afterschool Alliance
Peter Scales	Search Institute
Marc Schindler	Justice Policy Institute
Victor St. John	Child Trends
Noel Tieszen	Consultant in Public Policy and Nonprofit Management
Gina Warner	National AfterSchool Association
Roger Weissberg	University of Illinois Chicago

Appendix IV. Programs That Participated in the Study

Programs that participated in interviews and survey

- Administration for Children’s Services NYC
- Cheyenne River Youth Project
- Community Alternatives
- Connected Camps
- Creative Action
- Creative Solutions
- Department of Youth and Community Development
- Fiver
- Free Street Chicago Theater and Performance for Social Justice
- Generation Citizen
- Genesys Works
- Girls on the Run
- Hillside Work-Scholarship Connection
- Inside Out Literary Arts
- Kalamazoo Youth Development Network
- Mikva Challenge
- Mural Arts
- Northwestern Center for Development
- Our Piece of the Pie
- ourBRIDGE for Kids
- Playworks
- Project Morry
- Shakespeare Behind Bars
- Sherwood Forest
- Ucode
- Urban Alliance
- Wings for Kids
- Words Beat & Life

Programs that participated in survey only

- Boys and Girls Club of San Francisco
- Camp Fire USA
- Eastern Michigan University Bright Futures
- Ignite Afterschool
- Juma
- LA's BEST
- Outward Bound USA
- RYSE Youth Center
- Say Yes Buffalo
- West End House Boys and Girls Club, Boston

Appendix V. Data Collection Protocols

Documenting Outcomes and Measures in OST: Expert Interview Protocol

Introduction

Hi []. This is my colleague [Notetaker's name]. S/he'll be on taking notes or may jump in with a question or two. **Thanks for agreeing to** talk with us today.

Let's start with just a quick explanation of this project we're doing: The Wallace Foundation has engaged Child Trends to conduct a study to help inform the development of the Foundation's next initiatives related to out-of-school time (OST) programming for children and youth. More specifically, Wallace is interested in better understanding /what outcomes programs measure in their OST programming, what tools or methods (qualitative or quantitative) they use to measure outcomes (if any), and what barriers they experience when trying to collect high-quality outcomes data in OST programs. We expect today's discussion to take about 30 minutes. Do you have a hard stop at 1:30?

To gather information on intended outcomes, measurement tools, and gaps in measurement, Wallace recommended that we talk to strong programs to learn more about what they are doing. To help identify those programs, we developed a list of experts and are looking for recommendations you have of programs that would be good for us to interview.

These programs may or may not have been formally evaluated and they may or may not necessarily be well-known. It could be a large and well-known provider of services or a mid-sized program based in one locality.

Other things that might be useful for you to know is that Wallace is interested in programs that offer content in arts, civic engagement/social justice, career and workforce development, STEM, or general services, as well as programs focused on SEL and online programs.

In addition, they are interested in programs that serve school-age children and adolescents from marginalized communities and those that embed an equity lens in their work and are also hoping for us to talk to a subset of programs that serve systems-involved youth.

Can you please recommend 2-5 strong OST programs in the social and emotional learning sector for us to potentially interview?

We are also interested in gathering information from you about your thoughts about outcome measurement in OST.

I. Questions for Expert Informants

- Would you say that there are any common outcomes that most or many programs in [area] aim to target?
- What do you see is the main outcomes of focus for OST programs in [area]?
 - Probe on specific outcome areas
- Have you noticed any changes in outcomes of interest to OST programs/camps over time? Any changes in what programs/camps are interested in measuring? If so, why?
 - Probe on specific outcome areas
 - Are there any resources on OST measurement in virtual settings?

- Are there recent compilations of OST measures and tools that you're aware of? We are interested in global measures as well as those that might target a specific outcome or program type (STEM, SEL measures, arts, leadership, etc.)
- Do you find that most programs base outcomes measurements on a theory of change?
 - Probe
 - How many programs have a theory of change in the first place?
 - Go back to theory of change when analyzing outcomes?
- What would you say are the key gaps in measuring outcomes in OST?
 - Probe about the following:
 - Lack of available measurement instruments for certain outcomes
 - Weak measures
 - Lack of testing across populations
 - Ease of administration
 - Time
 - Financial resources
 - Usefulness to practice
- Are there other people you think we should reach out to speak about OST outcomes and measurement documentation?
- Are there any organizations/agencies that you think are thinking about outcomes and measurement in creative or innovative or particularly strong ways?
- Are there any additional resources we should review on outcomes, measurement instruments, or reports that we should review for our research?

II. Conclusion

- Is there anything I didn't ask about that you'd like to share about documenting and measuring outcomes in the OST field?

Thank you so much for taking the time to speak with us today, we really appreciate your insight.

Are there any tools for measuring outcomes that you or programs you've worked with have found useful? (Request a copy)

Documenting Outcomes and Measures in OST: Program Interview Protocol

Introduction

Hi, my name is [YOUR NAME] and this is [NOTETAKER'S NAME], we're from Child Trends. I will be leading our discussion today and [MY COLLEAGUE, name] will be taking notes. Thank you so much for taking the time to talk with us today. Let's start with just a quick explanation of this project we're doing: The Wallace Foundation has engaged Child Trends to conduct a study to help inform the development of the Foundation's next initiatives related to out-of-school time (OST) programming for children and youth. More specifically, Wallace is interested in better understanding what outcomes programs measure in their OST programming, what tools or methods (qualitative or quantitative) they use to measure outcomes (if any), and what barriers are in place to collecting high-quality outcomes data in OST programs. We expect today's discussion to take 30-45 min.

Your participation in the interview is voluntary. We're happy to skip any questions you prefer not to answer. What you say here will be kept private, which means that your name will not be shared or associated with what you tell us today without your permission. **[For programs only:** If it's okay with you, your organization will be included in a list of programs from which we gathered information.]

Also, as a sign of our appreciation for the time you will take to share your insights and experiences, we will send a \$50 gift card to you after the interview. This will be for a website called "Giftbit" and you'll be able to use it for a few different vendors such as Amazon, Target, or Starbucks amongst others. Is the email address we used to schedule the interview an ok one to send the link for this giftcard?

We plan to audio record today's conversation and may transcribe the recording to help us capture all the details from our discussion. The audio-recordings will be deleted once the study is complete. Is that ok with you, do we have your permission to record today's discussion?

Interview Questions

I. Interviewee Role and Background

Note for interviewer: The main goal of this section is to build rapport and gather background information on the interviewee and their organization.

To begin, we would like to hear a little about your role at the organization.

1. Can you tell us your title and a bit about your role at the organization?
 - a. [If unaddressed from the overarching question] **How long** have you worked at [name of organization]?

II. Background Information about the Program

2. Can you tell us a little about the OST programming offered at [name of program]?
 - a. What types of programming are offered in your organization?
 - b. What are the goals or focus of the programs? If there are too many programs to talk about, we can focus on 1-3?
 - c. [IF SURVEY DOESN'T HAVE INFO] What are the program locations? Note: this is geographic but also TYPE of place (i.e., school, rec center, etc.)

3. What is the target population of your program?
 - a. Can you tell me a bit more about the demographic backgrounds of the young people in your programs?
4. Do you consider your program or organization to have a strong focus on equity? If so, how does this focus show up?

III. Intended or Targeted Program Outcomes

5. What would you say are the intended or valued outcomes for participants in your program? (if they are having trouble because they have too many programs, you can ask for each program – if there are too many to talk about, work together to choose 1-3) → Probe: What does “success” look like for an individual in your program?
 - a. In what ways do you hope participants will benefit (in terms of outcomes) from your program?
 - b. How were these focal areas for outcomes identified?
 - c. How have your intended or target outcomes for program participants changed over time if at all?
6. Do you use a theory of change or logic model to guide outcomes measurement?
 - a. How do you use it?
 - b. If yes, would it be possible to share your theory of change or logic model with us?
7. In addition to what you have shared so far, are there any additional or secondary outcomes of your OST program(s)? These might be ones that you HOPE to see change in, but they are less of an intentional focus of the program or are more secondary?
8. Have you established shorter term (or interim) and longer-term outcomes that you hope your program will effect?
9. Has your program ever asked young people who participate in the program how they hope to benefit from their participation in the program?
 - a. Has youth provided input into the selection of program outcomes?
10. Do you track whether there are differences in outcomes across different groups or populations that you serve (race/ethnicity, age, gender, location, etc.)?

IV. Measurement of Outcomes

11. Have you gathered any evidence on what outcomes are resulting from your program? If yes, what have you gathered?
12. What type of data do you collect, if any, to measure whether participants are benefiting from the program?
 - a. **Probes:** Do you use surveys, interviews, administrative records, focus groups, observational assessments, tests, or other data collection instruments to assess whether participant outcomes are improving as intended?
 - b. Did you create the data collection tool(s) that you use to assess program outcomes? If not, was it an off-the-shelf tool or was it developed by an external evaluator?

13. Do you any other methods to document whether or how participants are benefiting from your program? Please describe what you use?
14. How did you decide what measurement tool to use? What other tools did you consider?
 - a. **[Probe if they only mention quant tools]** Do you only use quantitative measures (surveys, test scores) or do you also collect and look at qualitative measures? This could be focus groups or interviews or detailed observations? If yes, tell me more about that? Where? When? How?
15. [IF USING TOOLS] Would it be possible for us to have a copy of the tools you're using to collect data on participant outcomes?
16. Have you undergone evaluation by an outside, third-party evaluator?
 - a. If yes, what methodology was used to evaluate the program?
 - i. Probes: outcomes, pre-post, comparison group, implementation study, etc.
17. [OKAY TO SKIP] Once you analyze data and have information about outcomes, in what formats and with whom is this information shared? Anything else? Anyone else?
18. [OKAY TO SKIP] Are youth ever involved in your data collection, planning for data collection, data analysis or in interpretation of data?
 - a. If so, how?
19. [OKAY TO SKIP] Does equity show up at all in your measurement approaches/strategies? Do you disaggregate data by gender or by race/ethnicity? Or by other factors?

V. Gaps and Challenges in Measurement

Now we want to switch gears and ask for your perspective on measurement gaps, or challenges you have experienced as you collect data on participant outcomes.

20. Are there outcomes for program participants you want to measure that you haven't been able to?
 - a. PROBE: Why haven't you been able to measure them? Have you not been able to find a good instrument to assess a specific outcome of interest? Or you don't know what question to ask?
21. Are there any outcomes of interest that you've tried to measure but have struggled to do so in some way? In what ways have you struggled (e.g., people skip a specific question, poor measures, everyone gives the same response, etc.)?
 - a. Is there anything you've tried to measure that didn't work to measure for some reason? (e.g., lack of available measures, the best way to measure the concept is not clear, or the measures you've found don't perfectly align with your program model)?
22. What are the limitations of the tools or strategies that you've used or tried to use?
 - a. Probe: Do you feel like your ability to measure outcomes of interest has been hindered by lack of available tools, or by poorly designed tools?
23. Do you feel you have the staff capacity and/or other needed capacity to conduct the measurement you want to conduct?
 - a. If not, which is lacking?
 - b. Why (i.e., is it hard to hire? Hard to fund? Hard to make time for?)?

24. Have there been challenges or other issues with the measurement instruments or strategies you've used?
25. [OKAY TO SKIP] Are there any other tools for measuring outcomes that you or programs you've worked with have found useful? Please describe.
26. [OKAY TO SKIP] Aside from measuring participant outcomes, what other data does your program collect to assess its performance? Probes: implementation, quality, program fidelity, organizational capacity, staff satisfaction, parent/participant satisfaction, etc.
27. What kinds of support would improve the measurement of outcomes in your program?

VI. Conclusion

28. Is there anything I didn't ask about that you'd like to share about documenting and measuring outcomes in your program? Anything I should have asked about?

As a reminder, we would love if you could send us the [logic models, tools, etc.] we discussed today.

Thank you so much for taking the time to speak with us today, we really appreciate your insight.

Appendix VI. Survey

Program Outcomes Email Questionnaire

- Overview

The Wallace Foundation has engaged Child Trends to help inform the development of initiatives in learning and enrichment, in youth arts, and in the OST field at large.

Specifically, Wallace is interested in documenting and measuring outcomes in OST programs with a focus on arts, civic engagement & social justice, career & workforce development, and social and emotional learning. This includes online programs, programs that operate afterschool or during the summer, programs that serve systems-involved youth, and programs that embed an equity lens in their work. In addition, Wallace is interested in outcomes measurement in OST programs that promote positive identity formation, and programs that serve school-age children and adolescents from marginalized communities.

To learn about intended outcomes, the breadth of measurement approaches and instruments, and measurement gaps and challenges, we are reaching out to programs that were recommended by OST field leaders. We would like to know whether and how your program measures outcomes for program participants. We want to learn about the measurement tools you're using – or are interested in using – to assess participant outcomes, and any challenges you're experiencing pertaining to outcome measurement.

This information will inform work that Wallace does in the future to support OST programs across the country.

Thank you!

- Organization Name
- Name of program mentioned in the email, if different than organization name
- Thinking of the program mentioned in the email, how many sites does your program have?
 - 1
 - 2
 - 3
 - 4
 - 5
 - More than 5
- Who operates your program?
 - Government entity
 - Private community organization
 - Other

- Program Location (Select all that apply)
 - Parks and Rec
 - Schools
 - Community-based organization
 - Other

Program 1 (pick your largest or most important program)

- Program Type (Select all that apply)
 - SEL
 - Arts
 - Civic engagement
 - Social justice
 - Leadership development
 - Identity-focused
 - Career/workforce
 - Sports/physical fitness/health
 - Stem
 - Other
- When is this program offered (select all that apply)?
 - Before school
 - After school
 - Summer
 - Other
- Is this an online program?
 - Yes, my program was designed to be an online program
 - Yes, we've adapted because of COVID-19
 - No
 - Not sure
- Age Groups Served (select all that apply)
 - 5-12
 - 13-17
 - 18-20

- Do you consider your program or organization to have a strong focus on equity?
 - Yes
 - No
 - Not Sure
- What are intended outcomes for your program?
- Have you undergone a formal evaluation?
 - Yes
 - Date of most recent evaluation
 - Who conducted the evaluation?
 - What were the results of the evaluation?
 - What methods were used to evaluate your program?
 - No
- Do you have a logic model?
 - Yes
 - No
- Do you have a theory of change?
 - Yes
 - No
 - Do you have a performance management system (Efforts to Outcomes - ETO, Apricot, KidTrax, CitySpan, etc.)?
 - Yes
 - No
- Do you have a staff person who helps support data collection and/or analysis?
 - Yes
 - No
- Program Measurements
 - Do you use any tools to measure outcomes for your program's participants? These could include, but are not limited to, quantitative tools such as survey instruments or assessment measures, as well as qualitative tools such as interview protocols or focus group guides, or other data collection tools that you're using.
 - Yes
 - No

- If yes, what measurement tools does your program use (e.g., surveys, interviews, assessments, portfolio, performance)?
- What is the most challenging aspect of collecting data on youth outcomes for your program?
 - Are there any other youth outcomes you are unable to collect?
- What are the youth outcomes you're unable to collect and why aren't you able to collect them?
- What supports would be most useful to your program to improve your outcome measurement strategies?

If you have additional programs and would like to add background information for those programs, please select "submit" and then start over.

OR, press "finish" to submit your answers.

References

- Bronte-Tinkew, J., Moore, K. A., & Shwalb, R. (2006). Measuring outcomes for children and youth in out-of-school time programs: Moving beyond measuring academics. *Child Trends*.
<https://www.childtrends.org/wp-content/uploads/2013/08/2006-14MeasuringOutcomes.pdf>
- Cava Tadik, Y., Smith, E., Yu, D., Leathers, M., and Farris, J. (2019). Using connected technologies in a continuous quality improvement approach in after-school settings: The pax good behavior game. *Journal of Technology in Human Services*. 37. 1-22. 10.1080/15228835.2019.1588190.
- Hair, E. C., Moore, K. A., Hunter, D., & Williams Kaye, J. (2001). Youth development outcomes compendium. *Child Trends*. <https://www.childtrends.org/wp-content/uploads/2001/10/2001-13YouthOutcomesCompendium.pdf>
- Lantos, H., Redd, Z., Stratford, B., & Joshi, A. (2021). Demystifying data: Strategies and tools for making data more meaningful in OST programs. In *Measure, Use, Improve! Data Use in Out-of-School Time*. Charlotte, NC: Information Age Publishing. <https://www.infoagepub.com/products/Measure-Use-Improve>
- McCombs, J.S., Whitaker, A., & Yoo, P. (2017). *The value of out-of-school time programs*. RAND Corporation. <https://www.wallacefoundation.org/knowledge-center/pages/the-value-of-out-of-school-time-programs.aspx>
- McCombs, J.S., Augustine, C. H., Unlu, F., Ziol-Guest, K. M., et al.(2019). *Investing in successful summer programs: A review of evidence under the Every Student Succeeds Act*. RAND Corporation. <https://www.wallacefoundation.org/knowledge-center/pages/investing-in-successful-summer-programs-essa-review.aspx>
- Wilson-Ahlstrom, A., Yohalem, N., DuBois, D., Adler, P., & Hallaker, D. (2014). *From soft skills to hard data: Measuring youth program outcomes*. The Forum for Youth Investment. http://www.cypq.org/content/soft_skills_hard_data