


Documenting **Progress** and Demonstrating **Results**: Evaluating Local Out-of-School Time Programs


SEPTEMBER 2002

Priscilla Little
Harvard Family Research Project

Sharon DuPree
Consultant

Sharon Deich
The Finance Project

Harvard Family Research Project

The FINANCE PROJECT

About This Brief

This brief, Documenting Progress and Demonstrating Results: Evaluating Out-of-School Time Programs, is a collaborative publication between Harvard Family Research Project and The Finance Project. This brief is third in Harvard Family Research Project's (HFRP) series of briefs, entitled *Issues and Opportunities in Out-of-School Time Evaluation*, and part of The Finance Project's *Out-of-School Time* series. Through the publication of this brief, our organizations seek to provide out-of-school time programs with the evaluation resources necessary to improve their programs and to demonstrate results for sustainability.

About Harvard Family Research Project's Series of Briefs, Issues and Opportunities in Out-of-School Time Evaluation

The *Issues and Opportunities in Out-of-School Time Evaluation* briefs, funded by the Charles Stewart Mott Foundation, are user-friendly, practical documents that highlight current research and evaluation work in the out-of-school time field. These documents draw on HFRP's out-of-school time research in order to provide practitioners, funders, evaluators, and policy makers with information to help inform their work.

About Harvard Family Research Project

Founded in 1983 by Dr. Heather Weiss, Harvard Family Research Project conducts research about programs and policies that serve children and families throughout the United States. Publishing and disseminating its research widely, HFRP plays a vital role in examining and encouraging programs and policies that enable families and communities to help children reach their potential.

Harvard Family Research Project
Harvard Graduate School of Education
3 Garden Street
Cambridge, MA 02138
Tel: 617-495-9108
Fax: 617-495-8594
Email: hfrp@gse.harvard.edu
Website: www.hfrp.org

About The Finance Project's Out-of-School Time Technical Assistance Project

This tool is a part of a series of technical assistance resources on financing and sustaining out-of-school time and community school initiatives developed by The Finance Project with support from the Wallace-Reader's Digest Funds and the Charles Stewart Mott Foundation. These tools and resources are intended to assist policy makers, program developers and community leaders in developing financing and sustainability strategies to support effective out-of-school time and community school initiatives. For additional information on this project, see www.financeproject.org/osthome.htm.

About The Finance Project

The Finance Project is a non-profit policy research, technical assistance and information organization that was created to help improve outcomes for children, families, and communities nationwide. Its mission is to support decision making that produces and sustains good results for children, families, and communities by developing and disseminating information, knowledge, tools and technical assistance for improved polices, programs, and financing strategies. Since its inception in 1994, The Finance Project has become an unparalleled resource on issues and strategies related to the financing of education and other supports and services for children, families, and community development.

The Finance Project
1401 New York Avenue, Suite 800
Washington, DC 20005
202-587-1000 (phone)
202-628-4205 (fax)
http://www.financeproject.org

Acknowledgments

This evaluation guide represents a joint collaboration between Harvard Family Research Project and The Finance Project. We both acknowledge the support of An-Me Chung at the Charles Stewart Mott Foundation, as well as the valuable feedback provided by Jen Rinehart, 21st Century Community Learning Centers Program Officer; Cheryl Hayes and Heather Padgette, The Finance Project; and Heather Weiss and Julia Coffman, Harvard Family Research Project. Finally, the authors wish to acknowledge previous work on this topic conducted by Pam Miller, former researcher at Harvard Family Research Project, and Sara Watson, formally of The Finance Project.

Documenting **Progress** and Demonstrating **Results:** Evaluating **Local Out-of-School** Time Programs

September 2002

by Priscilla Little, Harvard Family Research Project,
Sharon DuPree, Consultant
and Sharon Deich, The Finance Project

Harvard Family Research Project

3 Garden Street
Cambridge, MA 02138
617-495-9108
(fax) 617-495-8594
www.hfrp.org

The FINANCE PROJECT

1401 New York Avenue, NW, Suite 800
Washington, DC 20005
202-587-1000
(fax) 202-628-4205
www.financeproject.org

Table of Contents

Introduction

Why Evaluate Out-of-School Time Programs?

For many community-based organizations, out-of-school time (OST) programming is uncharted territory. Because of its newness, relatively little is known about OST best practices, program implementation, cost-effectiveness, and impact. However, in these times of decreasing public resources and increasing and competing demands for public investments, funders, policymakers, and their constituents want to know which investments are effective and how programs can be improved. This situation makes it imperative that leaders implementing OST programs also implement systems and methods to document their progress and demonstrate results. In other words, programs need to begin to grapple with the complex issues of *evaluation*.

Across the country, there is a growing core of "flagship" out-of-school time programs, i.e., large-scale out-of-school time programs and initiatives that include rigorous evaluations that will support a data-driven case for the development and sustainability of quality out-of-school time programming. However, it is unrealistic to think that the "fleet"—i.e., most small-scale out-of-school time programs that are not part of larger initiatives—has the resources or expertise to conduct this type of evaluation. Yet, in this age of accountability, all programs, not just flagship ones, are being asked to conduct some form of evaluation to help guide their program and demonstrate results.

Given this context, it is imperative that "fleet" programs, including local 21st CCLC grantees, have the tools and resources necessary to conduct low-cost, high-yield evaluations. The purpose of this guide is to enable fleet programs to do just that.

The Purpose of This Guide

The purpose of this guide is to provide fleet-level practitioners with techniques, tools, and strategies they can use to improve their programs and track their effectiveness over time. It is also intended to provide information about multiple aspects of evaluation, so program stakeholders can become more informed participants and, thereby, more empowered in the evaluation of their programs.

NEW 21ST CENTURY COMMUNITY LEARNING CENTERS GUIDELINES

According to new 21st Century Community Learning Centers (21st CCLC) guidelines:

Each [local] grantee must undergo a periodic evaluation to assess its progress toward achieving its goal of providing high-quality opportunities for academic enrichment. The evaluation must be based on the factors included in the *principles of effectiveness*. The results of the evaluation must be: 1) used to refine, improve, and strengthen the program and to refine the performance measures; and 2) made available to the public upon request. (p. 33)

To meet the *principles of effectiveness* described in the law, programs or activities must be based on:

- An assessment of objective data regarding the need for before and after school programs (including summer school programs) and activities in schools and communities
- An established set of performance measures aimed at ensuring high-quality academic enrichment opportunities
- If appropriate, scientifically based research that provides evidence that the program or activity will help students meet the state and local academic achievement standards

Simply put, local 21st CCLC grantees must conduct program evaluation for the dual purposes of improving their programs and providing evidence of effectiveness on academic measures.

U.S. Department of Education. (2002, May) *21st Century Community Learning Centers non-regulatory draft guidance*. Washington, DC: Author. Available on the Internet at: www.ed.gov/21stcclc/draft_nrg_may.doc.

WHY CONDUCT PROGRAM EVALUATIONS?

In general, OST programs conduct an evaluation for one or more of three basic reasons:

- *To make management decisions.* Evaluation data help make internal management decisions about what is working and what is not working, where improvement is needed, and how scarce resources can be allocated.
- *To demonstrate accountability.* Evaluation data can be used to demonstrate to current funders that their investment is yielding the intended results. Evaluation results can also be used in marketing tools, such as brochures or published reports, that help promote a program to potential participants, the media, and potential funders.
- *To build a case for sustainability.* Evaluation results can be used to show the impact a program has had on an individual, family, school, or neighborhood, and thereby secure funding for sustainability.

The audience for this guide comprises executive directors of OST programs and administrators and staff of nonprofit organizations and schools that provide out-of-school time services. Secondary audiences include: intermediaries working with OST programs to evaluate program practices or designing assessments/evaluations for the sites with which they work; funders; and public policymakers. The funders and policymakers may be trying to understand and use local evaluation efforts when creating evaluation designs for national and statewide implementation studies, or they may be trying to understand the long-term impacts of state or federal programs that benefit children and families.

This guide is organized into four parts: Part I reviews the key issues of conducting program evaluation, Part II describes how a logic model can be a useful tool for program planning and evaluation, Part III describes the five-tiered approach to program evaluation and its applicability to OST evaluation, and Part IV provides practical suggestions on how to maximize the power and utility of evaluation results. Throughout the guide, "Evaluation Tip" boxes provide practical advice on conducting various aspects of program evaluation, as well as many illustrative examples from OST programs. Appendices include a glossary of evaluation terminology and a review of data collection methods. A selected list of out-of-school time evaluation resources accompanies this guide.

Evaluation Principles

There is no definitive answer to the question of how to evaluate a given program; however, some basic evaluation principles can anchor whatever evaluation approach is taken.

• *All programs can do at least some sort of evaluation.* Whether collecting attendance data, administering participant satisfaction surveys, or collecting youth outcomes data, all programs are capable of some rudimentary form of data collection that can feed into program evaluation.

• *The degree to which program services are established is important; program "age and stage" help determine how to select an evaluation plan.* An evaluation approach is dependent on the level of service development the program has achieved. Young programs are focused more on start-up concerns, such as recruitment, implementation, and staff training. Older and more stable programs can begin assessing the effects of their services. An evaluation that takes program evolution into account can plan ahead for future needs.

• *The input of all stakeholders is critical when planning an evaluation.* Local OST stakeholders can include parents, youth, program staff, school staff, health care professionals, community-based organizations, city government officials, police officers, and members of the parks and recreation department. Recognizing the stakeholders' importance to the evaluation process right from the outset increases their willingness to participate and helps avoid resentment and address concerns as they arise.

• *An evaluation can be an empowering process for all stakeholders*. Evaluation can lead to a feeling of empowerment. It can make people feel a part of the project-that what they do or say matters. By empowering stakeholders, programs can become stronger and more viable. By continuously assessing both service delivery and family progress, a program can generate valuable, ongoing feedback that can be used to develop better programs.

SAVE THE CHILDREN, in collaboration with the Aguirre Group, chose a participatory/empowerment approach to evaluate their program, **WEB OF SUPPORT**, because this approach promotes self-sufficiency within the organization and among associated community partners, allows for customized evaluations of a diverse selection of programs and activities, and supports efforts for continuous program improvement.

The participatory/empowerment approach also allows community partners to have ownership and involvement at all levels of the evaluation process. They are able to develop their own definitions of positive change and improvement, based on the populations they serve and needs within the community, identifying the stakeholders and their roles in evaluation design and implementation. This approach also allows the contributions and viewpoints of individuals at all levels of Web of Support to be acknowledged.

For more information on the Save the Children Web of Support evaluation, see the Harvard Family Research Project Out-of-School Time Evaluation Database. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/evaldatabase.html.

A Note on Clarifying Language

When it comes to the language of measuring results, a clear and consistent lexicon does not exist. As stakeholders begin discussions about evaluation, it is likely that a variety of terms will come into play. For the purposes of this guide, the following terms will be used:

- **RESULT:** A "result" is a bottom-line condition of well-being for children, families, or communities. It is a broadly defined, fundamental condition that government and citizens consider essential. One such bottom-line expectation of the community might be that all of its children should be born healthy. Results are umbrella statements that capture the comprehensive set of needs that must be met to achieve success. By definition, achieving these basic conditions of success requires concerted action by all sectors of the community.

  Some states and communities use the term "outcome" instead of "result." The meaning is the same. However, the term "result" avoids potential confusion with unrelated debates about outcomes-based education.

- **INDICATORS:** Indicators are measures, for which we have data, that gauge community-level progress toward agreed-on results. Because results are broad statements of what communities want for their children, no single indicator is likely to signal full attainment of any given result. Communities must decide what constellation of indicators add up to progress on each result and then require a community-wide, cross-agency effort. Indicators that are most often used by community-based organizations and schools include measurement of the following: decision making, civic responsibility or youth leadership skills, academic achievement, and gainful employment.

- **PERFORMANCE MEASURES:** Performance measures reflect the achievement of agencies and specific programs. As such, they gauge progress at the agency level rather than at the community level. Appropriate performance measures are closely related to an agency's mission and purpose and are within its ability to control. They are narrow measures of how well programs operate with their service populations, as part of a larger strategy to achieve results for the whole population. Examples of performance measures include improvement in attendance for youth attending after school programs or high school graduation rates for program participants.

Definitions of Key Evaluation Terms

TERM	DEFINITION	EXAMPLES
RESULT OR OUTCOME: (Community-wide) A condition of well-being for children, adults, families, or communities.	The desired result is the vision or goal of the initiative. Results generally cannot be measured directly but, rather, are composites of various indicators.	• Children ready for kindergarten • Children and youth succeeding in school • Children in nurturing environments
INDICATOR: (Community-wide) A measure, for which data are available, that helps quantify the achievement of the desired result for community-wide populations.	Indicators can: • Change over a few months or many years • Reflect changes in people or systems • Be expressed generally or as a specific level of achievement expected by a specific time • Be specified as rates or numbers	• Rates of children reading by the end of first grade • Rates of children left alone after school before an appropriate age
PERFORMANCE MEASURE: (Program-specific) A measure, for which data are available, of the impact of agency or program service delivery, or of the quantity or quality of agency or program service delivery. It quantifies the contributions of individual people or programs toward movement on the indicators.	• Performance measures reflect the contributions of individual stakeholders, workers, programs, or organizations. Performance measures can measure either "effect/product" or "effort/process."	
	Effect/Product measures can be: • Changes in knowledge, skills, attitudes, behavior, or status of the target population • Expressed generally or as a specific level of achievement by a specific time • Specified as rates or numbers	• Rates of children completing homework on time • Rates of child care teachers using new skills
	Effort/Process measures can include: • "Raw materials" that go into a program—dollars, in-kind donations • Levels of activity—number of classes, products • Measures of the quality of a service—customer satisfaction, error rates	• Number of students in classes • Hours of staff training • Student-teacher ratio

Watson, S. (2000). *Using Results to Improve the Lives of Children and Families: A Guide for Public-Private Child Care Partnerships*. The Finance Project and Child Care Partnership Project. Available on the Internet at: nccic.org/ccpartnerships/results.pdf (Acrobat file).

Part I: Key Aspects of an Evaluation—Getting Started

Once a program has identified and convened a group of evaluation stakeholders, it must address a series of questions to help plan for its program evaluation. This section of the guide examines eight questions that are essential to consider during the planning phase of program evaluation:

1. What is the program's stage of development?
2. What is the program trying to do?
3. What information do funders expect?
4. Who is the program's target population?
5. What does the program already know?
6. What does the program want to know?
7. What resources must be in place?
8. Who will conduct the evaluation?

Only after these questions have been addressed can a program choose an evaluation approach and begin to conduct an evaluation. Although answers to these questions are presented in a linear fashion, the questions and their associated tasks are interrelated and are often dealt with simultaneously during evaluation stakeholder meetings.

1. What Is the Program's Stage of Development?

For a program to attend to families and the community effectively and to become sustainable, evaluations should take the stage of the program into account (see box, "The Four Stages of Program Development"). The evaluation approach and the program aspects to focus on are functions of whether a program is a single-service, multiservice, coordinated-service, or collaborative program. For example, a relatively new program is not in a position to assess results, regardless of its level of service coordination. Following are some examples of evaluation choices that can be made at different stages of program development.

A single-service, stage one program, with the goals of improving academic performance, increasing social skills, and improving opportunities for parents to become involved in the program, could engage in basic data collection that would provide information about the program

THE FOUR STAGES OF PROGRAM DEVELOPMENT

In thinking about evaluation, program evolution can be viewed as a continuum of four stages:

STAGE ONE: This is the stage of the single organization, which may be a multicomponent program with multiple sites. At this stage of development, the evaluation might be an intra-agency assessment and include descriptive information about a program. Programs monitor service delivery and assess client satisfaction. For newer programs, the focus is on implementation of program components; more established programs can examine whether services reach those intended to receive them and the effects of these services. Many small-scale OST programs can best be described as stage one programs, even though they may have been in operation for several years.

STAGE TWO: This stage includes programs with formalized interagency linkages and at least some common planning procedures. In addition to the evaluation tasks carried out by stage one programs, evaluation plans in this stage include assessments of coordinated activities. Here programs might track participants across agencies, explore how the referral system is working, and assess the level of integration of coordinated services.

STAGE THREE: Programs in this stage share funds and comprehensive service networks and collaborate on staffing. In addition to the evaluation tasks carried forward from the first two stages, programs in the third stage measure their network of coordinated services and assess the implementation of their shared information system.

STAGE FOUR: This is the phase of true collaboration and occurs at the systems level; here community-level issues are examined. Programs are one component of a community effort to change the way services are delivered to children and families. At this point, programs examine the benefits of genuine systems change by assessing the linkage between agencies, determining any changes measured at the community level, and describing governance structures.

The **MAKING THE MOST OF OUT-OF-SCHOOL TIME** (MOST) initiative is a system-building initiative that seeks to contribute to the supply, accessibility, affordability, and quality of out-of-school time programs, especially for low-income children, by strengthening the system of out-of-school time programs. The MOST initiative involves three cities: Boston, Chicago, and Seattle. Under the umbrella of one or more lead agencies, each MOST city brings together the stakeholders in the system of out-of-school time programs to set priorities, do joint planning, share information, coordinate activities, and begin to develop citywide strategies for addressing after school challenges. The purposes of the MOST evaluation were to describe the MOST strategies in each city and examine their effectiveness. Additionally, the evaluation sought to describe the important dimensions of the system of out-of-school time programs in each city (supply, program quality, and financing) in order to provide context for understanding MOST efforts and to aid future investment in the field.

For more information on the MOST evaluation, see the Harvard Family Research Project Out-of-School Time Evaluation Database. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/evaldatabase.html.

and the participants it serves. For example, if the program operates by providing homework support and supervised free play during nonschool hours, evaluators can:

- Collect demographic data on program participants
- Document services provided and staff training
- Survey parents and participants for satisfaction

In contrast, a stage four program that truly collaborates at the systems level can evaluate systems change in addition to program effects. Evaluators of a multisite, school-based program offering comprehensive services through linkages to local- and state-sponsored out-of-school time programs can:

- Document program services and describe their operation
- Examine the evidence of genuine systems change and the effects of these changes, particularly at the community level
- Measure the degree and success of interagency coordination
- Assess parent, family, and child outcomes
- Examine advocacy activities

Although some community-based out-of-school time programs can be described as stage four programs because of their partnerships and program offerings, many small-scale OST programs can best be described as stage one programs even if they have been in operation for several years. This means that as single-site programs approach the task of evaluation, they need to be realistic about expectations for assessing program impact.

Community-based OST programs that are just starting the evaluation process will most likely conduct basic evaluation activities, such as documentation of program services and collection of demographic data. It is unrealistic to expect standalone programs in their first few years of operation to assess program or community impact. (However, as discussed below, it is possible to use the first few years of operation to lay the groundwork for this type of evaluation in the future.) In the case of funder-mandated evaluations, it is especially important to work with the funder to set realistic expectations for evaluation, taking the program's developmental stage into account.

2. What Is the Program Trying to Do?—Defining Goals

One of the initial steps in any evaluation is to define program goals and how services aim to meet them. Convening all out-of-school time stakeholders for the program and clarifying goals together helps all stakeholders, including staff, specify program content and intentions.

A useful approach to goal setting is the development of a logic model. A logic model is a concise way to show how a program is designed and can make a difference for a program's participants and community. On one sheet of paper, a logic model summarizes the key elements of a program, reveals the rationale behind the program's service delivery approach, articulates the intended results of the program and how they can be measured, and shows the cause-and-effect relationships between the program and its intended results. Part II of this guide explores logic model development and its application to out-of-school time program evaluation more closely. Whether or not a program chooses to develop a logic model, it is crucial that it has clearly articulated goals and objectives in order to determine the purpose of its program evaluation.

3. What Information Do Funders Expect?

Many programs enter the evaluation process with predetermined evaluation requirements. Funders, such as the U.S. Department of Education, may require programs to collect specific statistics about participants, their families, and the services they receive. In addition to these requirements, funders sometimes impose timeframes for the reporting of evaluation data, as well as the format and dissemination procedures for reporting results. Given these constraints, some strategies can help programs negotiate with funders about evaluation requirements:

- Work with funders to clarify what information they expect and when they expect it. Keep up an ongoing dialogue about the kinds of information funders are interested in and find useful.
- Try to allow for startup time, when needs assessment data and client characteristic data can be collected.
- Think collaboratively and creatively about effective evaluation strategies, including collecting data for program improvement.
- Be explicit about reasonable timeframes; for instance, it is unrealistic to expect information on long-term participant outcomes after only one year of program participation.
- Finally, try to negotiate separate funds for the evaluation component of the program. Be sure to note any additional staffing and resources needed for evaluation.

4. Who Is the Program's Target Population?

Out-of-school time programs serve a range of children and youth. Some focus on a specific population in the community, such as youth at risk of leaving school; others, such as 21st Century Community Learning Centers, serve children in high-poverty areas. Some OST programs provide activities for children alone, whereas community-based out-of-school time programs, such as the Beacons in New York City, offer a range of programming to meet the entire family's needs. It is important to document the intended target population, so that when the program begins to collect data on who it actually serves, it can compare intent to actuality.

5. What Does the Program Already Know?

Most programs start the evaluation process knowing more than they thought they knew about the program and its participants. Some have conducted previous evaluations and can draw on that data to craft new questions and new evaluation approaches. Programs that have previously conducted a needs assessment will already have data about community needs and the services currently being offered. Thus, current evaluation efforts can build on the knowledge gained from previous undertakings.

In addition to existing evaluation information, programs can turn to pre-existing data sets, such as Kids Count[1] and the U.S. Census Bureau, to obtain pertinent OST context data to help craft a rationale for program sustainability and expansion.

> ***Evaluation Tip***
>
> **Programs that have computer systems may have access to a wealth of information. A thoughtfully designed management information system (MIS) can assist in program operation and assessment. The main component of a MIS is a database for collecting, storing, and retrieving client and program information. Use of databases is especially important to collaborative efforts, due to the volume and diversity of the information collected. Building MIS systems into program proposals is a good way to ensure that a program is able to collect necessary evaluation data.**

6. What Does the Program Want to Know?

It is critical for programs to decide what aspects of their programming they want evaluation to inform. Answering

DETERMINING THE RIGHT EVALUATION QUESTION(S)

In order to:	Ask:
Respond to participant needs	• How can we better serve our families? • How do they view our program?
Respond to key stakeholders	• How do community members perceive our program? • Why did our public relations efforts work?
Understand community resources and the need for additional services	• Is our community satisfied with our efforts? • In what ways can our program affect our community?
Understand the services you provide and who you provide them to	• Who are we serving and how often? • What are the most useful amounts and types of service?
Improve on what you are doing	• In what ways is our program tailored to unique family and youth circumstances? • Are families and youth satisfied with the services we provide?
Describe and measure the impact of your services	• Are certain types and amounts of services more useful than others for our youth and our community? • Are we providing services that will help sustain our program over time?
Determine the costs of your program	• Is our program effective in attaining its stated goals? • How much do our programs cost, and are we providing services efficiently? • Is there a way to market our successful strategies to ensure future financial support?
Plan for the future	• What services can we add to better serve our families? • How can we coordinate these efforts?

the questions below can help form evaluation questions that are realistic and reasonable, given the program's mission and goals.

- What is the impetus for the evaluation? Does the program want basic facts about who it is serving? Is the program required to complete a more extensive evaluation for funders and the community?
- Is the program required to follow a specific evaluation report format? How much time does a program have to provide an evaluation report?
- How can the program generate support for the evaluation from all stakeholders? How will stakeholders be involved?
- How can programs generate information that is useful?

A review of several OST evaluations reveals that programs want to know the answers to the following general questions and look to evaluation to provide the answers:[2]

- Is my after school program responding to participant needs?
- Who is participating in my after school program?
- What is the staffing for my program? What training do they have and need?
- What after school resources are available in my community? What are the service gaps?
- What services does my after school program provide?
- How can I improve my after school services? Are participants, families, and the community satisfied?
- What is the impact of my after school program-academic achievement (e.g., increased school attendance and homework completion), youth development (e.g., increased positive school behavior), workforce development (e.g., employment preparation and support for families transitioning from welfare to work), prevention (e.g., reduced drug and alcohol abuse)?
- What are the costs of my after school program?

7. What Resources Must Be in Place?

Allocating resources to and funding an evaluation are key issues that programs need to address in order to make evaluation a reality. Each of these issues is discussed below.

ALLOCATING RESOURCES

Considering evaluation options involves assessing tradeoffs between what the program needs to know and the resources available to find the answers. Resources include finances, time, training commitments, and staff allocation, as well as technological aids such as personal computers and MIS systems. Programs can always start small and tailor evaluations to the number of people available to conduct the evaluation activities, the funds available, and the all-important time commitment. In fact, sometimes a small evaluation has the benefit of providing feedback quickly, thus enabling programs to make necessary changes in a timely fashion. As programs expand, assessment activities can also be expanded as people, training, and time allow.

PROGRAM PARTICIPANTS ARE A VALUABLE EVALUATION RESOURCE

Because young people are usually thought of only as individuals *receiving* services, we often forget that many youth are also capable of *giving* something back to a program or organization. Activities such as youth leadership programs, parent and youth advisory councils, and mentorships give both youth and parents the opportunity to participate in programs in ways that help them and others reach their personal and professional goals. Consequently, involving both youth and adults in the evaluation process can be very beneficial and make the end product richer and more useful to the organization. In small initiatives, particularly, youth and adults can serve as valuable resources.

Youth and adults can be involved in all aspects of evaluation design and implementation. Both youth and adults can provide input on what to look for and why. They can also help collect data. One youth organization used young people to help evaluate a program activity they were involved in. Another program had an outside evaluator train a young person to operate a computer-based data-collection system for the youth organization. That young person is responsible for organizing data on attendance and levels of participation in various program activities, information which is used in the organization's progress reports to funders.

ENGAGING FUNDING PARTNERS FOR EVALUATION

OST programs can turn to a variety of outside sources to support their evaluation efforts.

- A youth development program received resources from a local business that was interested in their job-training initiative. The business was interested in knowing if the program could provide trained entry-level staff.
- Sometimes local school districts have provided resources to after school programs that include literacy activities, because those types of programs, if effective, could enhance the academic knowledge of the students who attend their school.
- The Youth Development Institute, an intermediary organization that provides technical assistance and support for the New York City Beacons Initiative, raised funds from private foundations to do a year-long evaluation and an outcome study of that particular youth development model.

In planning an evaluation, programs often grapple with several frequently voiced resource barriers:

- Time—We already have enough to do.
- Money—We cannot afford to hire anyone, we cannot afford to pay our people more to complete additional tasks, and we cannot afford a computer to analyze all this.
- Expertise—Where do we begin and where do we go once we have our information? If we hire an outsider, will we lose control of our evaluation?
- Funders and Community—How do we know what they need?
- Consequences—In what way are we going to intrude on our families, and how will this affect them? Will our families tolerate these intrusions? Will the staff find these intrusions justified or acceptable?
- Validity—What constraints do our evaluation resources impose on the scale and type of evaluations that are possible?

FUNDING AN EVALUATION

Resources for evaluation, including funds for staff release time or external expertise, should be incorporated into proposals for all programs of an organization. It is easier to request resources and funding for an evaluation if a program has a clear idea of what it would like to measure, why it chose particular methods of measurement, and the ways in which it plans to monitor progress. Even if a funder does not agree to fund the evaluation process, the process of requesting funding for evaluation enables programs to determine what resources they will need to conduct an evaluation.

Evaluation Tip

What can influence evaluation costs?

- **Data-collection methods**
- **Number of sites included**
- **Length of evaluation**
- **Use of an outside evaluator**
- **Availability of pre-existing data sets**
- **Type of reports generated**

Many organizations can cover the cost of evaluation with resources from the ongoing budgets of their programs. However, when program resources cannot support evaluations, organizations have to find creative ways to obtain the resources needed to design and implement measures to track the effectiveness of their work. It is often necessary to pool resources from different entities to collect information about a variety of programs that could be used for different types of evaluations. Resources for evaluation can come from a range of interested entities in the community and can include cash, professional expertise, and staff time. Universities interested in research make great partners by providing expertise around design issues and assisting with data collection and analysis.

8. Who Will Conduct the Evaluation?

Program evaluators—the people with the skill and experience needed to implement an evaluation—are a vital resource. Sometimes staff and other OST program stakeholders may have the skills and experience necessary to design and implement an evaluation. Staff may have relevant pre-service coursework in evaluation, businesses are often eager to provide volunteers, and students at local colleges and universities are often eager to gain valuable on-the-job training by assisting in an assessment. At times, however, programs need the design and analysis expertise of an outside consultant. Further, some funders subtly suggest all facets of the evaluation be completed by an "objective" outsider, while other funders explicitly demand this.

There are advantages and disadvantages to consider when hiring an outside evaluator (See box below).

Evaluation Tip

Effective communication is integral to a successful evaluation. Whether the evaluation is carried out using in-house staff or outside assistance, someone in the program needs to act as the link between the evaluation team and those who request the evaluation.

HIRING AN OUTSIDE EVALUATOR

ADVANTAGES

- Can actually be cost-effective if an in-house evaluator would extend the process because of inefficiency or inadequate time and expertise
- Can provide a fresh perspective with new ideas
- Can focus attention solely on the evaluation
- Loyalty lies with the process, not particular people in the organization
- Potentially better at conflict resolution, due to outsider perspective

DISADVANTAGES

- Can be costly, especially if terms of the initial arrangement are renegotiated
- Might need to bring in someone from an entirely different locale
- May take a long time for the person to get to know the program and its people
- May inadvertently focus on issues that are not essential to the evaluation
- Absence of loyalty to the program can create conflict
- May lack necessary trust to keep lines of communication open and effective
- May lack familiarity with the families served or the community

ONE WAY TO CONSIDER THE OPTIONS OF USING AN OUTSIDE CONSULTANT IS IN TERMS OF AVAILABLE RESOURCES.

- Does the OST program have access to knowledgeable evaluators?
- How much time can program staff allocate to evaluation responsibilities?
- Does the program have the staff resources by which to develop in-house forms for evaluation purposes?
- What type of expertise will the evaluation really need? Can the program document client contact? Can the program conduct its own interviews?

ONCE PROGRAMS DETERMINE THE AVAILABLE RESOURCES, THEY CAN CONSIDER HIRING AN OUTSIDE EVALUATOR, WHO CAN BE EXPECTED TO:

- Observe the normal day-to-day activities of your program at length
- Be sensitive to the needs of participants, both families and staff
- Be willing to communicate, and do so in a way that is easily understandable
- Inspire change and assist in assessing processes to implement change
- Be able to help you determine what it is you need to know, rather than dictate your needs
- Promote site ownership of the evaluation

HIRING AN EVALUATION CONSULTANT

If a program does decide to obtain support from an external source, it will hire a consultant in the majority of situations. Often there are numerous consultants available to choose from, and it can be difficult to know where to start looking. Here are a few tips:

- Start by researching individuals recommended by people you know.
- Look for individuals who give free consultations, who have a reputation for the type of evaluation you want to do, and who are known to be flexible enough to work within your timeframe and your budget.
- Always ask for a proposal and a budget to help you make your decision and to be clear about what the consultant will do.

The following are some questions to ask evaluator candidates before selecting the evaluator who best matches your program's needs:

- How long have you been doing this work?
- What type of evaluations do you specialize in?
- What other organizations have you worked for?
- Can you show me samples of final evaluation reports?
- How much on-site training do you do for staff and administration?
- Have you ever worked with children? Youth? Families?
- Do you conduct telephone consultations? If so, are extra costs involved?
- Do you conduct follow-up visits? If so, are extra costs involved?
- Will you be the individual we are working with or do you have partners?
- What is your method of communication regarding challenges, progress, and procedures?
- Can we review and respond to reports?

Part II: Logic Models

In designing an evaluation, it is important to be clear about the goals of the program and to be realistic about the expectations for which it can be held accountable. In Part I, Section 2, we noted that a logic model is a useful tool to assist programs in defining their goals and figuring out the focus of their evaluation. Many organizations rely on logic models to help guide planning and program improvement efforts. This part of the guide describes how to develop a logic model.

Since creating a logic model requires a step-by-step articulation of the initial goals of the program and the proposed activities that will be conducted to carry out those goals, it also helps a program chart the resources it needs to carry out the evaluation process.

Information for logic models can come from a variety of sources. The **PROJECT FOR NEIGHBORHOOD AFTERCARE** in Nashville, Tennessee, interviewed program staff to learn what the goals of the program were, how the program intended to achieve these specific goals, and what program actions were considered key to achieving the goals of the program. The resulting program logic models or flow charts were then presented individually to the staff for confirmation that the logic models accurately reflected how the program was intended to make a difference in the lives of students. The program's goals and logic models are expected to evolve with the program.

For more information on the Project for Neighborhood Aftercare evaluation, see the Harvard Family Research Project Out-of-School Time Evaluation Database. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/evaldatabase.html.

Why Develop a Logic Model?

There are many different logic models, just as there are many frameworks for learning about results-based decision making. Programs may want to consider several approaches before choosing or modifying a particular model to fit their circumstances. Once the concepts are familiar, the same understanding may be applied to any logic model. The logic model, like an architect's scale model of a building, is not meant to be a detailed "blueprint" of what needs to happen. It is a place to lay out the major strategies, to see how they fit together and whether they can be expected to "add up" to the changes the OST stakeholders want to see.

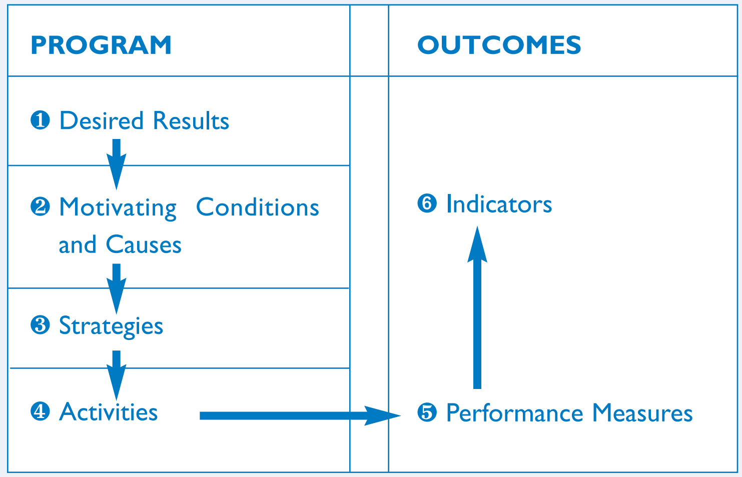
Creating a logic model at the beginning of the evaluation process not only helps program leaders think about how to conduct an evaluation, but also helps programs choose the programmatic aspects they want to evaluate. Practitioners who have created logic models for evaluations have been most successful when they have included administrators, staff, and, when possible, youth in the process. While a logic model is not the only tool needed to measure performance, it is one of the key methods used to assist organizations in tracking program progress.

What Are the Elements of the Logic Model?

Based on information from a guide published in 2000 by the Child Care Partnership Project, a logic model has several elements.[3]

- *Result:* The logic model starts with the program's desired result-what is its vision, what does it ultimately want to happen or exist in its community?
- *Conditions and Causes:* The next step is to examine the "story" behind the results-what are conditions, causes, circumstances, or factors that need to change in order to move toward the result? They are what cause the result to happen or prevent it from happening. These conditions should be based on research or, at least, specific experience about what works.
- *Strategies:* The next step is to examine the broad approaches or strategies that will affect those conditions or causes in ways that help the organization reach the desired results. Both the strategies and the activities need to include all partners who will work to implement the plan.

DEVELOPING A LOGIC MODEL



Developing a logic model essentially requires filling out its elements (represented by the boxes to the right) with details based on your OST program. While there is no "correct" order in which to do this, it is suggested that you start with the left column of the logic model, and move counterclockwise.

❶ Decide on your desired result.
Decide what result (goal or vision) you are trying to achieve. This result should express your vision for children, youth, and families.

❷ Decide on the major conditions and causes behind that result.
Starting with the desired result, discuss the indicator data, research, and experience that suggest the major conditions, causes, and factors that affect your result. Think broadly before narrowing your scope.

❸ Develop strategies.
Strategies are broad approaches to child, youth, and family well-being. Since you may not be able to implement all of the desired strategies at once, it may be helpful to prioritize them based on your list of indicators. Once you have prioritized your strategies, you can develop a more detailed action plan and performance measures for those activities you plan to implement first.

❹ Develop activities.
Activities are the individual services or interventions your program uses to implement your strategies. Ask yourself: On a day-to-day basis, what does the staff in my organization do? What services do we provide? Work plans may be useful for identifying this list (e.g., homework help and tutoring, mentoring, rap sessions). Then develop an initial list of activities to implement the strategy.

❺ Develop performance measures.
Considering the activities you plan to pursue, develop some performance measures that can track progress on the activities. These measures include measures of effort (amount of activities you are implementing), as well as measures of effect (the changes in the child, the youth, or the family population you are serving).

❻ Decide on the major indicators you want to affect.
Using the indicators from step two, choose a small number of indicators that you want to affect. Remember, indicators reflect community-wide changes in children, youth, and families, so you will need partners to change these numbers.

(See page 22 to continue the logic model steps)

ELEMENTS OF THE MODEL "THE PROGRAM"

Desired Results

Improve the physical, social, and emotional well-being of children.

Improve children's academic development and performance.

Motivating Conditions and Causes

- Many parents working outside of the home
- Children with unstructured and unsupervised time in the after school hours
- Low academic performance among low-income children
- Lack of positive adult-youth relationships
- Children at greater risk for involvement in crime and substance abuse in the hours after school
- Television as the most common activity for children after school

AS MEASURED BY "THE OUTCOMES"

Indicators

Ultimate Indicators

- Reduced rates of substance use among teens
- Reduced teen pregnancy rates
- Reduced numbers of violent acts among adolescents and teens
- Reduced dropout rates
- Increased percentage of students graduating from high school
- Increased percentage of students attending college

Require community-wide effort to affect

Interim Indicators

- Improved test scores in reading, math, or science
- Reduced numbers of antisocial behaviors or behavior problems
- Decreased student suspensions
- Improved grades

ELEMENTS OF THE MODEL "THE PROGRAM"

OST Program Strategies

- Youth development and leadership
- Academic enrichment
- Curriculum development and enrichment
- Collaboration

OST Program Activities

- Homework help and tutoring
- Mentoring
- Rap sessions
- Arts activities
- Recreation activities
- Technology training
- Literacy activities
- Career counseling and development
- Community service or work projects
- Intergenerational activities
- Conflict resolution training

AS MEASURED BY "THE OUTCOMES"

Program Performance Measures

Measures of Effect

- Development of emotionally supportive relationships with adults
- Increased emotional adjustment
- Increased social competence
- Higher self-esteem and confidence
- Improved study habits
- Higher numbers of honors or awards received
- Improved peer relationships
- Improved attitudes toward school
- Improved school attendance and decreased truancy

Can be affected by the OST program

Measures of Effort

- Number of children served in the OST program and participant demographics
- Number of classes, sessions, or trainings held
- Number and type of products developed
- Measures of program cost-effectiveness
- Parent and child satisfaction rates with the OST program

- *Activities:* The final step identifies the activities, programs, interventions, and services that need to be implemented to carry out the strategies.

Each activity the program pursues may change only one or a few of the conditions and causes that would need to change in order to reach the desired result. Each activity may also affect only a subset of the total target population. So the program leaders need to think about how much activity program staff need to implement to bring about the desired improvements. Each element should also be based on research or experience that explains why a bad condition exists or a good condition does not exist (the conditions and causes) and what will fix it (strategies and activities).

On the previous page is an example of a logic model for an out-of-school time program.[4] Following this example is a matrix of sample performance measures and possible data sources.

What Performance Measures Do Out-of-School Time Programs Use?

Once a program has created and articulated its desired results, motivating conditions and causes, its OST strategy, and its program activities, the next step in logic model development is to determine how to use performance measures to assess progress. As previously stated, an indicator is a measure, for which there are data, that gauges community-level progress toward agreed-on results. Because results are broad statements of what communities want for their children, no single indicator is likely to signal full attainment of any given result. Further, it is unrealistic to expect that a single OST program can “move the needle” on these community-wide indicators. Rather, programs can affect performance measures that track progress on program activities and purposes. Choosing performance measures that are reflective of program activities ensures natural methods of measurement of short-term performance and more accurately shows clear correlations between the goals of the program and the results.

Two families of performance measures are familiar to most OST practitioners. One type involves academic achievement measures. These measures assess individual performance related to school achievement, such as report grades or scores on standardized tests. The other type involves youth development measures, such as the measurement of decision-making skills or civic responsibility. Youth development measures are related to the less tangible effects a program has on an individual, but have the potential of making a long-term impact on an individual’s academic achievements.

Compiled from a review of city-level OST evaluations, the table on page 21 illustrates the range of performance measures OST programs use to measure academic achievement and youth development, as well as the data sources used to obtain information about the measures.[5] It clearly shows that there are many ways to define and measure academic achievement and youth development. The selection of measures to use should be tied to the goals of the program and the activities it provides.

Measures and Data Sources for Outcome Areas

OUTCOME AREA	DATA SOURCE
Academic Achievement	
• Academic performance in general	• Parent report, principal report
• Attendance or absenteeism	• School records, parent report, principal report
• Attendance in school related to level of program participation	• School records
• Attendance in school related to achievement	• School records, standardized tests
• Attitude toward school	• Child report
• Behavior in school*	• Standardized behavior scales by teachers
• Child's ability to get along with others	• Parent report
• Child's liking for school	• Parent report
• Child's communication skills	• Parent report
• Child's overall happiness	• Parent report
• Cooperation in school	• Child report
• Effectiveness of school overall	• Principal report
• Effort grades	• School records
• English language development	• Child report
• Expectations of achievement and success	• Child report, teacher report
• Family involvement in school events	• Principal report
• Grade point average	• School records
• Grades in content areas (e.g., math, reading)	• School records, parent report
• Homework performance	• Parent report, principal report
• Learning skills development	• Teacher report
• Liking school more	• Child report
• Motivation to learn	• Parent report, teacher report
• Reading	• Child report, principal report, test scores
• Safety—viewing school as a safe place	• Child report
• Scholastic achievement assessed by knowledge about specific subjects	• Parent report
• Standardized test scores	• SAT-9, state assessments
Youth Development	
• Adults in the OST program care about youth	• Child report
• Awareness of community resources	• Child report
• Behavior change toward new program component	• Parent and child report
• Child's self-confidence	• Parent report
• Exposure to new activities	• Principal report
• Facing issues outside of OST program	• Child report
• Interaction with other students in OST	• Child report
• Interest in nonacademic subjects (e.g., art, music)	• Child report
• Leadership development and opportunities	• Child report
• Opportunities to volunteer	• Child report
• Productive use of leisure time	• Child report
• Sense of belonging	• Child report
• Sense of community	• Child report
• Sense of safety	• Child report
• Sources of support for youth	• Child report

Measures and data sources compiled from a review of findings from 13 city-level evaluations; for brevity, "child" refers to youth of any age participating in the OST program.

* School behaviors included in the scales are frustration, tolerance, distraction, ability to ignoring teasing, nervousness, sadness, aggression, acting out, shyness, and anxiety.

HOW DO PROGRAMS USE LOGIC MODELS FOR EVALUATION AND LEARNING?

Logic models help OST programs describe their activities and identify outcomes and measures that will help them assess progress and results (i.e., steps one through six on page 18). However, they can also help programs move forward with evaluation in terms of putting plans in place to collect data on the identified measures and to use that data and the logic model for learning. Four additional elements (steps seven through ten) that can be added to the logic model toward that end are identified below.[6] The figure on the next page illustrates the relationships between those four elements and the six elements of the logic model discussed earlier.

❼ Data Sources and Methods

These are the sources for the data needed to track indicators and performance measures. Ask yourself: Now that I have identified my measures, how will I get the needed data in the most resource-efficient way? If you used the criterion that data should already be available for the indicators you have chosen, then you should already know their data sources and how often they are available. However, you also need to determine how often to report that information, how to deliver it, and who will get it. Performance measures will likely require additional data collection that either you or your evaluator will conduct. Some of that information, such as the measures of effort, you can probably track on your own. However, you may need to use an external evaluator to collect data on the measures of effect (e.g., sources: standardized testing, state or local government databases; methods: surveys, focus groups, interviews).

❽ Evaluation Questions

These are the questions you want to have answered by the data or decisions that you want to make based on your data. You should be able to make decisions based on your indicators and performance measures. Ask yourself: What strategic decisions can I make based on the information that will be generated? What consequences should be tied to achievement for good or bad performance? (e.g., Are the indicators showing progress and, if not, does that mean the OST program needs to be modified?)

❾ Stakeholders

These are the individuals who have a vested interest in the OST program. They need to know the answers to the evaluation questions and to be involved in learning from the data being collected. Ask yourself: Who is interested in or will benefit from knowing my program's progress on its indicators and performance measures (e.g., board members, funders, collaborators, program participants, community members, and other individuals or organizations)?

❿ Mechanisms for Learning

These are the periodic or regular opportunities that exist for pulling together and reporting the data being collected and for bringing together stakeholders to learn from and make decisions based on the data. Ask yourself: What opportunities exist or need to be created to focus on and learn from the evaluation (e.g., staff, stakeholder, or board meetings; regular evaluation reports; strategic retreats)?

Coffman, J. (2002). *Learning from logic models in out-of-school time.* Cambridge, MA: Harvard Family Research Project. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/resources.html.

Logic Models for Learning and Evaluation

PROGRAM		OUTCOMES	EVALUATION AND LEARNING
Desired Results			**❼ Data Sources and Methods** Where the data needed to track the indicators and performance measures will come from.
Motivating Conditions and Causes	↔	**Indicators**	**❽ Evaluation Questions** Questions, based on indicator data and movement, that determine whether the strategy needs to be modified. Questions, based on the performance measures, that determine whether the project is working as intended, what lessons have been learned, and how the project may need to be modified in order to achieve better results.
Strategies			**❾ Stakeholders** The funders, collaborators, participants, and other individuals or organizations with a vested interest in the program who need to be involved in learning from the data being collected.
Activities	↔	**Performance Measures**	**❿ Mechanisms for Learning** The opportunities for stakeholders to come together to learn from and make decisions based on the data about the program.

Coffman, J. (2002). *Learning from logic models in out-of-school time*. Cambridge, MA: Harvard Family Research Project. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/resources.html.

Part III: Choosing an Evaluation Starting Point—The Five-Tiered Approach

Once a program has decided what to evaluate, determined the key players to be involved, and developed a logic model to identify goals, objectives, strategies, and performance measures, the next step is to think about how to collect and measure the needed information. While there are dozens of approaches to program evaluation, the five-tiered approach[7] is particularly well-suited for local OST programs that are interested in using evaluation for program improvement. This approach is nimble enough to allow for program growth and change, which is an integral part of the emerging OST field.[8]

WHY THE FIVE-TIERED APPROACH FOR OST EVALUATION?

The five-tiered approach is well-suited to OST program evaluation for three important reasons.

- All programs are able to do at least some evaluation-tier one, pre-implementation planning, is something that every program can and should do.
- A program's information needs evolve as it develops and, therefore, its evaluation approach must be nimble enough to accommodate program change.
- Evaluation is an ongoing, cyclical process-feedback from one phase of the evaluation can be used to shape the next phase.

As its name suggests, the five-tiered approach to program evaluation describes the evaluation process as a series of five steps. The step on which a program starts an evaluation is largely determined by the program's developmental stage. For example, a new program may want to begin by developing a logic model during a pre-implementation planning phase, while an older, more established program might focus on refining its services and disseminating program results based on performance measurement data that it collects. The five steps are:

1. *Pre-implementation planning.* Identify the out-of-school time services already available in the community, as well as the services needed but not provided. Develop a logic model that reflects the goals, objectives, and strategies that a program will use to address the service gaps.
2. *Document program services.* Monitor the program by systematically documenting the out-of-school time services it provides; establish indicators of success and a set of benchmarks to record progress.
3. *Compare program intent to actuality.* Look at original goals and who the program is serving, and gauge whether the program is meeting benchmarks.
4. *Fine-tune program services and evaluation methods.* Discuss evaluation data with staff and brainstorm about program improvement; refine indicators and benchmarks; begin to document program results.
5. *Produce evidence of effectiveness.* Develop research design and evaluation measures; demonstrate your program's worth to participants, the community, and funders.

Tier One: Pre-Implementation Planning

The purposes of pre-implementation planning are to:

- Document the need for a particular program in the community
- Demonstrate the fit between community needs and the proposed program
- Provide "data groundwork" or baseline information
- Identify indicators of change that are of interest to various stakeholders

Integral to this tier are two important tasks: (1) developing a logic model (described above) and (2) conducting a needs assessment to better understand the "fit" between the program and the community.

"Need" can be defined as "the gap between the problem and the existing efforts, resources, and programs to deal with the need."[9] When assessing needs, a program must acknowledge the services already available and identify those that could be of potential help if provided to the youth and families participating in the OST program.

Summary of the Five-Tiered Approach to Program Evaluation

TIER	PURPOSE	TASKS
Tier One: Pre-Implementation Planning	• To document the need for a particular program in the community. • To demonstrate the fit between community needs and the proposed program. • To provide "data groundwork" or baseline information. • To identify indicators of change that are of interest to various stakeholders.	• Describe program vision, mission, objectives, goals, and characteristics. This step can involve the development of a logic model to articulate the program's theory of change. • Conduct community needs assessment to support establishment of the program. • Establish baseline status on key indicators.
Tier Two: Service Documentation	• To document and examine a program's use, entrenchment, or penetration in the target population. • To justify current expenditures. • To increase expenditures. • To build a constituency. • To begin to elicit and examine questions about early implementation.	• Accurately describe program participants and services provided; provide accurate cost information. • Build organizational capacity for self-assessment and program modification.
Tier Three: Program Clarification	• To provide information to program staff to improve the program. • To re-examine and modify, if necessary, the theory of change, and the short-term indicators and results. • To prepare, if resources permit, for tiers four and five.	• Question basic program implementation and outcome assumptions. • Examine implementation. • Clarify and restate program's mission, goals, objectives, strategies, and relationships with cooperating providers and the community. • Examine evidence of change on indicators. • Continue to build a culture of a learning organization.
Tier Four: Program Modification	• To modify the theory of change, if necessary, and to provide information to staff to improve the program. • To document and report changes in participant status, examine effectiveness, and build accountability.	• Examine progress on short-term outcomes. • Develop measurable indicators of success for outcomes. • Decide on data analysis procedures. • Assess differential effectiveness among individual clients. • Assess community views and program awareness.
Tier Five: Program Impact	• To potentially modify the theory of change and provide information to improve the program. • To document changes in participants, examine effectiveness, build accountability. • To model a learning organization and effective practice. • To contribute to knowledge development and theories of change in the substantive field of interest, or to the refinement of evaluation practices. • To document program effectiveness. • To produce evidence of differential effectiveness among program approaches. • To suggest program models worthy of replication.	• Develop and execute impact or outcome evaluation. • Delineate specific impact outcomes that are to be achieved. • Identify measures that can assess enduring or participant lifestyle changes. • Develop comparison standards (e.g., experimental or quasi-experimental design). • Conduct analysis and write up reports. • Work with stakeholders to develop reporting strategy for diverse audiences.

ASSESSING COMMUNITY NEEDS AND STRENGTHS

The Children's Aid Society in New York City identifies inclusion of a needs assessment as essential to program design:

To design a community school that will effectively respond to the needs and strengths of the community, there must be a clear understanding of what those specific needs are and what assets, resources, and services already exist in the community. This may seem like an obvious statement, but sometimes in the rush to provide services in communities where there is much work to be done, we fail to step back and take the time to examine the unique characteristics of the community. This can result in a needless duplication of services or, worse, the oversight of some basic and critical services needed.

The Children's Aid Society. (2001). *Building a Community School* (3rd ed.). New York: Author.

Needs assessment is important in building a data-driven case for program and service expansion. Too often a program provides youth and their families with the services that it can easily offer instead of the services they need.

Also called the "needs assessment tier," this first tier of program evaluation is well-suited to new programs that are trying to develop in ways that best fit the needs of the community. Tier one is also relevant, however, to programs that are already underway. Although successfully established programs have, by definition, already accomplished some of the tasks of this phase, even those that are up and running may want to take a step back and determine if their planned programs have adequately addressed the needs of their community. In so doing, they can raise awareness about a problem and garner additional support before attempting to alleviate it. Needs assessments look at potential participant recruitment, public awareness of specific social issues, and how the community perceives the program's goals and activities. They can also measure the sense of program ownership or "buy in" by the community.

Evaluation Tip

The following questions are important to consider prior to implementing a needs assessment:

- **Does the program have the time to commit to a thorough assessment?**
- **Can the program free up staff time for this effort?**
- **Have program stakeholders interacted positively with community leaders?**
- **What are the needs issues the program is most interested in?**
- **Who in the community will be able to inform the needs assessment?**
- **What information has already been gathered about serving children and families in the community?**
- **What specific program elements are planned?**
- **How much will it cost?**
- **What benefits can the participants and community expect?**
- **Can community needs be reassessed as the program evolves?**

Programs can use several methods to carry out a needs assessment:

- Resident surveys
- Interviews with OST stakeholders (community organizations and leaders, businesses, cultural organizations, and religious associations)
- Use of existing statistics and other information, such as census data or department of education information
- Surveys of other organizations serving youth
- Current literature review on the problem areas under discussion
- Newspapers and other media to help identify needs

Evaluation Tip

Be sure to obtain relevant information that others have collected. Community organizers may have completed some of the research you need about your community's strengths and capacities. Consider these sources of information when assessing your community's pre-existing services: local colleges and universities, public schools, hospitals, social service agencies, police and fire departments, libraries, and recreational resources.

Tier Two: Documenting Program Service

Monitoring a program by systematically documenting the service it provides enables programs to gather information that can be reported to multiple stakeholders. Many funding agencies (including the U.S. Department of Education, through the 21st CCLC grants program) require this type of data. Even when not required, however, the information is useful—programs can use the accumulated data to see if they are reaching their intended target population and to justify continued funding. Service documentation assists programs in understanding each of the service components and provides valuable information about resource allocation. Service documentation can also be the foundation on which later evaluations are built. Monitoring program services will help discover answers to critical questions:

- What services or activities does the program offer?
- Is the program serving the intended population of youth and their families?
- Are services differentiated among different populations?
- How are staff resources allocated?
- How are funds utilized?

Evaluation Tip

Regardless of the program's target population, a necessary component of an evaluation is determining who actually uses the services provided by the program. Essentially, programs need to answer the following set of questions:

- **Are some groups over-represented?**
- **Are others under-represented?**
- **Is the program serving people it never expected to attract?**
- **How are program resources allocated among different participant segments?**

Documenting program services has two basic purposes. One is to justify expenditures. The other is to describe the frequency, content, and quality of OST program activities, including:

- Activities offered
- Participation rates
- Staffing patterns
- Staff training
- Transportation usage

Evaluation Tip

When designing the system of service documentation, program leaders need to consider several questions:

- **Are we keeping too many records, thereby consuming valuable staff time?**
- **Are our forms designed to collect the relevant data efficiently?**
- **Do we actually use the information we collect?**
- **Can the information we collect be used for more than one purpose, thereby conserving both time and money?**
- **Are we doing too little documentation, so that we do not know if we are intervening as planned?**
- **Is our documentation helpful for planning?**
- **Are we expecting so much data that our reporters become careless?**

The most basic way to document program service is to:

- Devise a multipurpose intake form that describes participant characteristics
- Develop forms to record program activities and participation rates
- Keep accurate records of staff and their training

Evaluation Tip

In general, devising a workable intake form is an extended process. But functional forms facilitate the information gathering already required and result in additional data—at no extra cost. Keep the form as simple and complete as possible. You do not want to hamper either your staff or the families filling out the forms. Increased monitoring duties can be burdensome and may cause the staff to rebel. If you end up with missing data, you will incur higher evaluation costs in the long run. The challenge is to find that delicate balance between comprehensiveness and practicality.

The **PROJECT FOR NEIGHBORHOOD AFTERCARE** (PNA) is a school-based, no-fee, after school program in Nashville, Tennessee that serves the children in the neighborhood of each sponsoring school from 3pm to 6pm on school days. PNA offers academic, recreational, and cultural enrichment activities, while providing students with a safe, supervised environment during the after school hours.

PNA conducted an evaluation to examine several aspects of its program operation: participation characteristics and patterns, program activities, satisfaction, improvements in school performance, study habits, attendance, attitudes toward school, and suggestions for improving the program. The evaluation's primary focus was to provide feedback to the program management and funders in order to assist program development and successful implementation. In essence, PNA evaluators used tiers two and three of the five-tiered approach to document program implementation and use the information for program improvement.

For more information on the PNA evaluation, see the Harvard Family Research Project Out-of-School Time Evaluation Database. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/evaldatabase.html.

Tier Three: Program Clarification

This step of evaluation determines if the program is doing what it set out to do. This is the time for a program to examine its original goals and who it is serving, to gauge whether the program has been implemented in the way it was intended. The program's logic model can be used to compare program actuality to program intent. This step focuses on how program decisions are made and by whom. It describes what the program looks like in "real life" and analyzes how it operates on a day-to-day basis.

A program compares program intent to program actuality in order to:

- Examine service data to determine if targeted youth were appropriately defined.
- Review and revise program objectives.
- Provide timely and pertinent feedback to staff members.
- Make program adjustments and improvements.

In this tier, stakeholders reflect on some or all of the following questions:

- What were the program's intended activities?
- Were all activities implemented?
- Are services offered appropriate to targeted youth and their families?
- Are some youth excluded?
- What did participants think about these offerings?
- Can activities be observed in a minimally intrusive way?
- How will the feedback be used?
- How can the program do a better job of serving youth and their families?

There are many ways to compare what a program intended to provide with its actual services including:

- Compare service documentation data with logic model goals and objectives.
- Ask staff members to observe each other's activities.
- Ask staff members and participants to keep a journal of their experiences with the program.
- Enlist participating families to give you feedback in small group sessions.
- Develop a client satisfaction survey.

Evaluation Tip

Looking at the differences between what was planned and what happened prepares you to move beyond numbers and statistics to analyze participant and program data through extensive description. By combining qualitative and quantitative data, your evaluation will provide you with a well-rounded perspective on program results.

Tier Four: Program Modification

Having answered whether or not program services match intended program goals (tier three), programs can begin to fill in the gaps and fine-tune program offerings. At this phase of evaluation, programs can discuss evaluation data with the staff and other key stakeholders, brainstorm about program improvement, document program effectiveness, and begin to consider issues of evaluation information dissemination.

At this stage, stakeholders can:

- Revisit short-term goals and determine indicators of success
- Consider how data analyses can reinforce conclusions
- Review the standardized measures
- Measure community response to the program
- See if family and youth goals correspond to program goals

The **LA'S BETTER EDUCATED STUDENTS FOR TOMORROW** (LA's BEST) program is an after school program created in 1988 by the mayor's office as a partnership between the City of Los Angeles, the Los Angeles Unified School District, and the private sector. The program has five goals: (1) a safe environment, (2) enhanced opportunities through the integration of an educational support structure, (3) educational enrichment activities to supplement and deepen the regular program, (4) recreational activities, and (5) interpersonal skills and self-esteem development. LA's BEST students receive homework assistance, library activities, and a variety of other activities, provided in safe environments.

The program conducted a multiyear, multisite evaluation to: describe selected program characteristics, services, and operations of the LA's BEST program; describe the perceived effects of the program; and explore selected issues raised in earlier evaluation studies. It has used evaluation results for program enhancement and improvement and to generate concrete, useful recommendations for improving program quality. When evaluation results were released in June 2000, they revealed that LA's BEST supports gains on standardized tests, better school attendance, higher academic achievement, and English-as-a-second-language learning. The evaluation results led to Mayor James Hahn's 2001 expansion of the program to 101 school sites, as well as to the program's use as a model in cities across the country.

For more information on the LA's BEST evaluation, see the Harvard Family Research Project Out-of-School Time Evaluation Database. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/evaldatabase.html.

Evaluation Tip

When considering ideas for program and evaluation modification or refinement, consider the questions below:

- **How intrusive are our evaluation procedures?**
- **Can we be more sensitive to our families' needs while still fulfilling our evaluation needs?**
- **Should we consider hiring an outside expert to aid in evaluation?**
- **How do we package our success stories so people can absorb our findings?**
- **What have we accomplished?**

Tier Five: Program Impact

Evaluation Tip

Prior to starting tier five, it is important to consider the following question: Is the program sufficiently developed to begin this process? Designing an evaluation to verify program effectiveness takes the time, resources, and expertise that many small-scale programs do not have. Further, even with adequate resources, a new program should set realistic expectations about its ability to demonstrate program effectiveness in the first year of program operations.

After conducting tiers one through four, some program leaders are ready to tackle the complex tier of determining program effectiveness. Unlike some of the other program evaluation tiers, which can be conducted by staff and other program stakeholders, this step will most likely require the assistance of outside expertise. In this tier, programs need to address issues of evaluation design in order to build a credible case that the results observed are the consequence of program participation. Increasingly, "flagship" evaluations are using experimental and quasi-experimental designs to determine program effects; however, it is more realistic for the "fleet" to consider non-experimental designs that attempt to collect data on participant results without comparison or control groups (see Appendix A for more information on experimental designs).

Evaluation Tip

When choosing the research design that best suits your evaluation, consider these questions:

- **What is the most stringent research design that can be implemented to assess the multiple components of our program?**
- **Is it possible to implement an experimental research design with random assignments to treatment and control groups? Can we consider this design choice for at least a sample of our participants?**
- **If not, can we locate a satisfactory comparison group?**
- **If not, what other research designs can be implemented and still satisfy community and funder requirements?**
- **What kinds of assessments are appropriate for purposes of legislative oversight and accountability?**
- **What kinds of outcomes will prove compelling to public policymakers concerned with preventing or improving public programs?**
- **What types of outcomes will provide evidence useful for local program development?**
- **Can we develop a feasible set of cross-site measures?**

In this tier of program evaluation, programs examine some, or all, of the following questions:

- Does the program produce the intended results?
- Does it work better for some participants than others?
- How did the content of our services effect change?
- Do we see enhanced child or parent development, or both, as a result of our program?
- In what ways has the community benefited from our program?
- How can our findings influence policy decisions?

Programs collect formal evidence of program effectiveness for multiple reasons that meet stakeholder needs. Sometimes stakeholders, including funders, want evidence of the effects of the overall program. With increasing interest in evidence-based best practices, and in order to promote quality OST programming, many evaluators and researchers are interested in the effects of specific program components and their relationship to youth outcomes. Often funders and others are interested in outcomes data in order to identify program models for scale-up and expansion.

In 1995, the **EXTENDED-DAY TUTORING PROGRAM** was implemented with first through fourth graders in Title I Memphis City Schools. The program used a curriculum based on Success for All (SFA), with the purpose of improving participating students' literacy skills. A quasi-experimental evaluation was conducted to identify the strengths and weaknesses of the program, as well as to examine preliminary academic outcomes after the first year of implementation.

The evaluation employed a matched-pair design in which one student in each pair participated in the tutoring program while the other student did not.* Students in grades two through four were matched by the principal, using the 1995 Total Reading Normal Curve Equivalent (NCE) scores on the Tennessee Comprehensive Assessment Program (TCAP) as the primary criterion for matching and attitude and behavior as secondary matching criteria. There were 328 matched pairs included in the analysis. While first graders also participated in the tutoring program, they were not included in part of the evaluation since they do not undergo TCAP testing.

* Matched-pair design enables the evaluator to create comparison groups that are similar to treatment groups on the characteristics that are deemed to be most relevant to program outcomes. This allows for comparative statistical analyses of differences between the two groups.

For more information on the Extended Day Tutoring Program evaluation, see the Harvard Family Research Project Out-of-School Time Evaluation Database. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/evaldatabase.html.

21ST CENTURY COMMUNITY LEARNING CENTERS PROGRAM EVALUATION—AN EXAMPLE OF MULTIPLE TIERS OF EVALUATION

Local 21st Century Community Learning Center programs must indicate how they meet the *principles of effectiveness* described in the No Child Left Behind Act of 2001. According to the statute, local programs or activities must be based on:

- An assessment of objective data regarding the need for before and after school programs (including summer school programs) and activities in schools and communities
- An established set of performance measures aimed at ensuring high-quality academic enrichment opportunities
- If appropriate, scientifically based research that provides evidence that the program or activity will help students meet the state and local academic achievement standards

Simply put, new 21st CCLC programs will be asked to:

- Conduct a needs assessment (tier one of the five-tiered approach)
- Identify performance measures (tier four)
- Where appropriate, conduct scientifically based evaluation research (tier five)

States will also develop performance reporting mechanisms similar to the currently used Annual Performance Report, which collects data on service documentation (tier two) that can be used for program clarification and improvement (tier three).

Part IV: Getting the Word Out

It is important to consider key stakeholder audiences—program participants, families, staff, administrators, director and board, local policymakers, community leaders, and key media—when writing an evaluation report. This section of the guide offers some practical advice on how to write reports that will have maximum utility to program stakeholders.

How to Write Comprehensive Reports

A comprehensive evaluation report should include:

- An executive summary of findings
- A description of the program that was evaluated (including funders, objective, and other key information)
- If applicable, information about the overall organization and how the evaluated program fits into that organization's mission
- The purpose of the evaluation, the evaluation process, the methods used to carry out the process (tools, protocol, type of data-collection systems)
- Information (along with resumés or other biographical information) on the individuals who were responsible for carrying out the work (including the lead person and other members of the evaluation team)
- Information about the target area and statistical information about the target participants, and the connection between those general statistics and the purpose of your evaluation
- A report and discussion of findings: What did the evaluation reveal? How will the findings help facilitate program improvement? How can the information be used for program sustainability?

It is important to determine which parts of the comprehensive report are appropriate for which audiences. Also, think about how to share results with program participants, community partners, and all other OST stakeholders.

How to Report Findings

When reporting findings, keep in mind the various audiences that are interested in the outcome of the evaluation efforts. Some funders want extensive discussion of major and minor findings, while others are interested in only a summary of major findings. In either case, it is important to ensure the clarity of the presentation of findings by:

- Reporting findings one at a time
- Clearly discussing the performance measure related to the finding, the outcome of the information collected, how the finding relates to the overall goal of the program and, if available, other research in the field related to the same topic
- Thinking about the implications of findings for practice and possibly public policy

Some evaluations lend themselves to a discussion of recommendations that might assist other practitioners who want to design a similar program.

Be Creative When Presenting Data

In addition to including the information necessary to meet funder and other stakeholder expectations, the readability of the report can be enhanced by including illustrative examples that bring the data to life. During the course of the evaluation, programs usually collect data from multiple

BEYOND THE TRADITIONAL EVALUATION REPORT

Here are some ideas from programs that have tried to go beyond issuing the traditional report:

- Include testimonials from satisfied families.
- Use drawings and graphs to portray activities families engage in over time.
- Produce brief videos with the help of experienced community volunteers or local ad councils.
- Give out periodic newsletters to families and community members.
- Have a community volunteer photograph families and the staff and include these in your report.
- Develop press releases.
- Encourage media coverage of your success stories.

sources and in a variety of formats—not only facts and figures, but case examples and feedback from families. Incorporating this information into the final evaluation report can make the report more user-friendly to multiple stakeholder audiences. This is especially important if programs are trying to capture the attention of local media, who often like the "human interest" side of evaluation.

Evaluation Tip

Using both quantitative and qualitative data greatly increases evaluation options. More importantly, combining both types of data gives a program the ability to describe itself more fully. Facts, figures, and feelings make for a rich description of a program's families and services; do not emphasize one at the expense of another.

How to Communicate Results to Stakeholders

How do programs talk about or disseminate the results of their evaluation? The answer to this question connects back to accountability and the purpose of evaluation efforts. If a program conducts an evaluation for program improvement, results will most likely be communicated to administrators, staff, parents, and participants in various ways, many of them informal. However, if a program is conducting a formal evaluation for funders, it must consider more formal ways to communicate results, such as comprehensive reports or formal presentations. Organizations have disseminated the results of their evaluations through:

- Presentation of results at staff meetings for management and staff within the organization
- Presentations at luncheons or seminars for external parties who are stakeholders (such as collaborative partners)
- Presentations at regional meetings or national conferences in the fields of education, youth development, family strengthening, or public policy
- Comprehensive reports for partners, community businesses, or funders who are looking for concrete documentation of program impacts
- Executive summaries or full reports posted to program or organization websites

Sharing results does not have to mean sharing every bit of information that was collected. Deciding which information is the most important to which audience is the key to communicating evaluation findings.

Conclusion—Tips for Pulling it All Together

There is no one recipe for program evaluation. Out-of-school time programs use a variety of approaches, methods, and indicators-both to collect data for program improvement and to demonstrate the effectiveness of their programs. It is up to leaders in individual program sites to decide the best approach-one suited to the program's level of development, the needs of program participants and the community, and sponsor expectations. However, a final set of questions serves as a "litmus test" prior to the implementation of any evaluation.

Are all the necessary stakeholders included in the evaluation process? The success of an evaluation hinges on the participation of multiple OST stakeholders. Achieving buy-in and cooperation for the evaluation from the outset paves the way for a smoother evaluation process.

Is the evaluation purpose clear? Clarity about the purpose of the evaluation is key to active and accurate participation by all stakeholders. Generally, all evaluations are conducted for one or more of four main purposes:

- To understand both the resources in your community and the need for additional services
- To understand the services provided and to whom they are provided
- To improve the program
- To describe and measure the impact of program services

Do evaluation indicators map onto program goals? Integral to the success of an evaluation is matching evaluation indicators to program goals. An evaluation cannot measure and describe activities and services that programs do not provide. Developing a clear logic model can help ensure that evaluation indicators match program goals.

Is the information useful? An evaluation is not successful unless the information discovered is useful. Whether the evaluation documents participation, assesses program impact, provides feedback for program refinement, or monitors progress in meeting goals, the information should aid in decision making or suggest improvements.

Endnotes

1. KIDS COUNT, a project of the Annie E. Casey Foundation, is a national and state-by-state effort to track the status of children in the United States. The 2002 KIDS COUNT data book is available on the Internet at: www.aecf.org/kidscount.

2. Based on a review of 21 after school program evaluations in the Harvard Family Research Project Out-of-School Time Evaluation Database. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/evaldatabase.html.

3. Watson, S. (2000). *Using results to improve the lives of children and families: A guide for public-private child care partnerships*. The Finance Project and Child Care Partnership Project. [Available on the Internet at: nccic.org/ccpartnerships/results.pdf (Acrobat file).]

4. From Coffman, J. (2002). *Learning from logic models in out-of-school time*. Cambridge, MA: Harvard Family Research Project. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/resources.html.

5. For the complete review, please see Little, P., & Traub, F. (2002). *Evaluating municipal out-of-school time initiatives—Learning from the field.* Cambridge, MA: Harvard Family Research Project. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/resources.html.

6. From Coffman, J. (2002). *Learning from logic models in out-of-school time.* Cambridge, MA: Harvard Family Research Project. Available on the Internet at: www.gse.harvard.edu/hfrp/projects/afterschool/resources.html.

7. Jacobs, F. H. (1988). The five-tiered approach to evaluation: Context and implementation. In H. B. Weiss & F. H. Jacobs (Eds.), *Evaluating family programs* (pp. 37-68). Hawthorne, NY: Aldine de Gruyter.

8. For those familiar with the U.S. Department of Education's *Continuous Improvement Management Guide*, the five-tiered approach operates on similar assumptions about evaluation.

9. Linney, J. A., & Wandersman, A. (1991). *Prevention plus III: Assessing alcohol and other drug prevention programs at the school and community level.* Rockville, MD: U.S. Department of Health and Human Services.

Out-of-School Time Evaluation Resources

Resources Available From Harvard Family Research Project

Youth Involvement in Evaluation and Research. This is the first brief in this series, *Issues and Opportunities in Out-of-School Time Evaluation.* It draws on information collected from focus group interviews with representatives of 14 programs that are involving youth in their evaluation and research efforts. It examines the elements of successful youth-involved research projects and offers short profiles of the 14 organizations included in the study.

Evaluation of 21st Century Community Learning Center Programs: A Guide for State Education Agencies. The second brief in the *Issues and Opportunities in Out-of-School Time Evaluation* series, this brief offers an in-depth look at the new 21st Century Community Learning Center evaluation requirements (both performance measurement for accountability and program evaluation) and provides practical suggestions about how to implement state and local 21st CCLC evaluations. It includes a checklist of issues to consider when designing state and local 21st CCLC evaluations.

Learning from Logic Models in Out-of-School Time. A logic model can be a powerful tool for illustrating a program's theory of change to program staff, partners, funders, and evaluators. Moreover, a completed logic model provides a point of reference against which progress toward achievement of desired outcomes can be measured on an ongoing basis, both through performance measurement and evaluation. This brief offers an in-depth review of logic models and how to construct them.

HFRP Out-of-School Time Evaluation Database. This database of out-of-school time evaluation profiles provides information, in an accessible way, about evaluation work of both large and small out-of-school time programs and initiatives. Each profile contains an overview of the out-of-school time program or initiative that was evaluated, as well as detailed information about each evaluation report produced about that program. Electronic links to actual evaluation reports, where available, are also provided, as are contacts for program directors and evaluators.

The Evaluation Exchange. This is HFRP's evaluation periodical. The Fall 2000 and Spring 2001 issues highlighted emerging issues in the out-of-school time arena and featured articles by out-of-school time evaluators, researchers, and practitioners. Both issues included a "New and Noteworthy" guide to selected print- and Web-based after school resources. The fall issue also included an interview with leading out-of-school time researcher and advocate, Michelle Seligson; the spring issue included an interview with practitioner and author, Jane Quinn.

These, and other evaluation resources, are available free on the HFRP website at www.hfrp.org. They can also be ordered by contacting the publications office by mail at 3 Garden Street, Cambridge, MA, 02138, by telephone at 617-496-4303, or by email at hfrp_pubs@gse.harvard.edu.

Resources Available From The Finance Project

The Financing Strategies Series. This series focuses on strategies for financing children and family services, as well as community building and development. Each strategy brief presents options for generating resources or for using resources effectively to fund specific programmatic needs. Each also illustrates these options with examples drawn from initiatives around the country and highlight relevant considerations to assist decision makers who are weighing the alternatives. The first two briefs, *Thinking Broadly: Financing Strategies for Comprehensive Child and Family Initiatives* and *Sustaining Comprehensive Community Initiatives: Key Elements for Success*, provide policy leaders with clear guidance on how to find funding and sustain it over the life of the program.

Out-of-School Time Technical Assistance Project. This series of tools provides technical assistance resources on financing and sustaining out-of-school time and community school initiatives developed by The Finance Project. These tools and resources are intended to assist policy makers, program developers and community leaders in developing financing and sustainability strategies to support effective out-of-school time and community school initiatives. Selected titles from this series include: Using Title I to Finance Out-of-School Time And Community School Initiatives; Financing Transportation Services for Out-of-School Time and Community School Programs, and Using the Community Development Block Grant to Support Out-of-School Time and Community School Initiatives.

Other policy briefs specifically about evaluation include: *A Guide to Successful Public-Private Partnerships for Out-of-School Time and Community School Initiatives*. Sharon Deich wrote this guide which is designed to provide policy makers, program leaders, system-building advocates, and others with practical information on creating and maintaining public-private partnerships. *Using Results to Improve the Lives of Children and Families: A Guide For Public-Private Child Care Partnerships*. This guide provides a user-friendly context and framework for thinking about, measuring and reporting the results of public-private partnerships. The tool provides information to enable partnerships to plan and collect information to guide decision making and show their effects and effectiveness. It includes specific examples of indicators that address both the processes of building and sustaining partnerships and the results of partnership efforts.

The Out-of-School Time web page www.financeproject.org/osthome.htm. For a complete listing of all of The Finance Project's resources on Out-of-School Time, visit the Out-of-School Time web page. The web page includes links to tools and materials, profiles of successful programs, financing strategies as well as links to other resources.

Other Resources

A Field Guide to Outcome-Based Program Evaluation, The Evaluation Forum, www.evaluationforum.com

Beyond the Bell: A Toolkit for Creating Effective After-School Programs. North Central Regional Educational Laboratory. Available on the Internet at: www.ncrel.org/after/bellkit.htm.

Continuous Improvement Management Guide for 21st Century Community Learning Centers. Available on the Internet at: www.ed.gov/offices/OUS/PES/21cent/cim226.pdf (Acrobat file).

Measuring and Evaluating Child Program Outcomes. School Age Review, 2000. www.nsaca.org/publications.htm.

The Program Manager's Guide to Evaluation, Administration on Children, Youth and Families. Available on the Internet at www.acf.dhhs.gov/programs/core/pubs_reports/hs/hs.html

Use of Continuous Improvement and Evaluation in After-School Programs: Final Report 2001. Prepared by the Center for Applied Research and Educational Improvement, College of Education and Human Development, University of Minnesota. Available on the Internet at: education.umn.edu/CAREI/Programs/mott/default.htm.

Appendix

APPENDIX A: Evaluation Terminology

ACCOUNTABILITY

A public or private agency, such as a state education agency, that enters into a contractual agreement to perform a service, such as administer 21st CCLC programs, will be held answerable for performing according to agreed-on terms, within a specified time period, and with a stipulated use of resources and performance standards.

BENCHMARK

(1) An intermediate target to measure progress in a given period by using a certain indicator. (2) A reference point or standard against which to compare performance or achievements.

DATA-COLLECTION METHODS

Document Review. A review and analysis of existing program records and other information collected by the program. Information analyzed in a document review is not gathered for the purpose of the evaluation. Sources of information for document review include information on staff, budgets, rules and regulations, activities, schedules, attendance, meetings, recruitment, and annual reports.

Interviews or Focus Groups. Conducted with evaluation and program or initiative stakeholders, including staff, administrators, participants and their parents or families, funders, and community members. Can be conducted in person or over the phone. Questions posed are generally open-ended. The purpose of interviews and focus groups is to gather detailed descriptions, from a purposeful sample of stakeholders, of the program processes and the stakeholders' opinions of those processes.

Observation. An unobtrusive method for gathering information about how the program or initiative operates. Observations can be highly structured, with protocols for recording specific behaviors at specific times, or unstructured, taking a more casual "look-and-see" approach to understanding the day-to-day operation of the program. Data from observations are used to supplement interviews and surveys in order to complete the description of the program or initiative and to verify information gathered through other methods.

Secondary Source or Data Review. Sources include data collected for similar studies for comparison, large data sets such as the Longitudinal Study of American Youth, achievement data, court records, standardized test scores, and demographic data and trends. Data are not gathered with the purposes of the evaluation in mind; they are preexisting data that inform the evaluation.

Surveys or Questionnaires. Conducted with evaluation and program or initiative stakeholders. Usually use a highly structured interview process in which respondents are asked to choose from predetermined answers on the survey, administered on paper, through the mail, or, more recently, through email and on the Web. The purpose of surveys and questionnaires is to gather specific information from a large, representative sample.

Tests or Assessments. Data sources include standardized test scores, psychometric tests, and other assessments of the program and its participants. These data are collected with the purposes of the evaluation in mind.

EVALUATION DESIGN

Experimental Design. Experimental designs all share one distinctive element—random assignment to treatment and control groups. Experimental design is the strongest design choice when interested in establishing a cause-and-effect relationship. Experimental designs for evaluation prioritize the impartiality, accuracy, objectivity, and validity of the information generated. These studies look to make causal and generalizable statements about a population or about the impact of a program or initiative on a population.

Non-Experimental Design. Non-experimental studies use purposeful sampling techniques to get "information rich" cases. Types include case studies, data collection and reporting for accountability, participatory approaches, theory-based or theory-grounded approaches, ethnographic approaches, and mixed-method studies.

Quasi-Experimental Design. Most quasi-experimental designs are similar to experimental designs except that the subjects are not randomly assigned to either the experimental or the control group, or the researcher cannot control which group will get the treatment. Like the experimental designs, quasi-experimental designs for evaluation prioritize the impartiality, accuracy, objectivity, and validity of the information generated. These studies look to make causal and generalizable statements about a population or about the impact of a program or initiative on a population. Types include comparison group pretest/post-test design, time series and multiple time series designs, nonequivalent control group, and counterbalanced designs.

FORMATIVE OR PROCESS EVALUATION

Formative evaluations are conducted during program implementation in order to provide information that will strengthen or improve the program being studied—in this case, the after school program or initiative. Formative evaluation findings typically point to aspects of program implementation that can be improved for better results, such as how services are provided, how staff personnel are trained, or how leadership and staff decisions are made.

INDICATOR

An indicator provides evidence that a certain condition exists or certain results have or have not been achieved. Indicators enable decision makers to assess progress toward the achievement of intended outputs, outcomes, goals, and objectives.

PERFORMANCE MEASUREMENT (ALSO CALLED PERFORMANCE MONITORING)

According to the U.S. General Accounting Office, it is "the ongoing monitoring and reporting of program accomplishments, particularly progress toward pre-established goals" (sometimes also called outcomes). Performance measurement is typically used as a tool for accountability. Data for performance measurement are often tied to state indicators and are part of a larger statewide accountability system.

SUMMATIVE OR OUTCOME EVALUATION

Summative evaluations are conducted either during or at the end of a program's implementation. They determine whether a program's intended outcomes have been achieved. Summative evaluation findings typically judge the overall effectiveness or "worth" of a program, based on its success in achieving its outcomes, and are particularly important in determining whether a program should be continued.

APPENDIX B: A Review of Data-Collection Methods

SURVEYS

Surveys are tools designed to collect information from a large number of individuals in a specific time period; the data collected can be compiled into an easily read diagram or chart. Surveys are an excellent way to obtain background and participation data for your program. Demographic information, such as age or economic status of participants, is easily obtained by the survey method. Many agencies use initial survey forms to obtain information about the interests of a set of individuals in their program. In addition, surveys can be used to get some idea of the individual progress of participants, according to predetermined performance measures. As with all data collection, it is important to make sure that the type of information you can obtain from a survey is the type of information that will be useful to you. The benefits and drawbacks of using the survey method are listed below:

BENEFITS	DRAWBACKS
• Low cost • Less subjective • Easy to analyze • Can be repeated at regular intervals • Easy to administer	• Monolithic view of program • Too constrictive • Questions need to be correctly interpreted • Unable to capture variables that might affect outcomes • Dependent on participants' acceptance and completion

INTERVIEWS

Interviews allow programs to gather detailed information about participants, either one-on-one or in small group settings. Interviews require a protocol (or set of questions) designed to elicit specific information from participants. Interviews can be used to collect information about the quality of your program or to measure performance of individual staff or the inner group dynamics of specific activities. Interviews are most often used to collect subjective or qualitative data, such as how participants feel about a particular activity, or whether or not they found participating in a certain activity helpful to them in a specific way. Youth development programs are often interested in finding out whether or not attendance in a youth leadership group helped a young person make better decisions about activities related to pregnancy prevention, substance abuse prevention, or career development. An academic tutoring program would want to find out whether or not students feel the extra assistance improved their grades in a certain subject area. A family support program would want to know whether or not a workshop series on adolescent communication and development helped parents interact better with their teenagers. The benefits and drawbacks of interviews are as follows:

BENEFITS	DRAWBACKS
• Open-ended questions • More in-depth information • More diverse information • More accurate information	• Time consuming • Data are subjective • Data interpretation is more challenging • Must develop an effective system of recording information

CASE STUDIES

A case study focuses on one individual over a determined amount of time. Agencies can choose to follow a set number of individuals over the course of the program to get an in-depth picture of program impacts on those individuals. Participants can be chosen randomly or as a result of specific criteria. For example, an after school homework assistance program might decide to do case studies of a small number of participants in the program to obtain a deeper understanding of how homework assistance affects overall school performance for middle school students. A case study takes an intensive look at an individual's participation in a program and the affect on that individual's life. Case studies look at different aspects of individuals' participation in a program and may even include (secondarily) certain aspects of their home lives.

Case studies can include formal interviews, informal contacts such as telephone conversations or conversations in corridors, and observations. In most instances, because case studies are a very intensive method of obtaining information, program staff carry out this evaluation method. Program staff usually have already built a relationship with the individual in the study and also have several opportunities to interact. The case study method of evaluation also requires the development of tools to guide the relationship between the evaluator and the case study individual. The drawbacks and benefits of case studies are as follows:

BENEFITS	DRAWBACKS
• In-depth understanding • Inform statistical data collection • Effective way to learn about cultural backgrounds and customs of participants • Can provide more holistic view of participants	• Time consuming • Very limited sample, not generalized • Can become burdensome to staff, participants, or both • Data can be subjective

FOCUS GROUPS

Focus groups are often used to probe—to get information about a certain topic from a large group of individuals who may have a variety of different perspectives. The information from focus groups can be used to inform program development, or to find out what is going on in a particular field. For example, in a research study for the Annie E. Casey Foundation's Making Connections Initiative, focus groups were conducted for parents, adults, and youth workers to determine family strengthening strategies in youth development organizations. The information from the focus group (along with interviews and research) helped document family strengthening activities in youth organizations. Focus groups follow a structured protocol. The benefits and drawbacks of focus groups are as follows:

BENEFITS	DRAWBACKS
• Obtain information from a large number • Less time consuming than interviews or case studies • Less staff involvement data • Less invasive to participants than case studies • Informs evaluation focus	• Difficult to record data from numerous individuals at one time • General consensus as opposed to in-depth data • Challenging to analyze • Need skilled facilitator to conduct • Requires large physical space to conduct

OBSERVATIONS

Some evaluations rely on observations. Observation is not looking for solutions to problems or for answers to certain questions, but is merely noticing what is going on around you. To use this method of evaluation, know that you need a person to do this task whose only job is to observe for a particular period of time. It is important to engage staff in self-reflection before they begin to observe in order to identify individual biases and how those biases or insights strengthen or weaken the observer's position. The benefits and drawbacks of observations are as follows:

BENEFITS	DRAWBACKS
• Assist staff in gaining perspective • Inform evaluation focus • Enhance other types of data • Less time consuming than certain other methods	• Less in-depth than other methods • Requires distancing from activities and participants • Can feel intrusive

LONG-TERM TRACKING

Long-term tracking requires following an individual over an extended period of time (such as 10 years) to determine whether or not the individual is following a predetermined pattern of behavior that was anticipated as a result of a previous intervention. This method of evaluation is usually chosen by organizations that are conducting a large-scale, multi-method evaluation. Outside consultants or research organizations are usually hired to design the evaluation and to collect and analyze data. The benefits and drawbacks of long-term tracking are as follows:

BENEFITS	DRAWBACKS
• A record of program impact • Multi-method data collection • Assistance from external experts • Organized data management systems	• Costly • Difficult to track participants long-term • High staff and management involvement • Can be invasive to programming