CHAPTER 1

FOUNDATIONS OF PROGRAM EVALUATION

After reading the introduction you should be able to

1. Provide a basic definition of program evaluation
2. Understand the different activities conducted by a program evaluator
3. Understanding the difference between formative and summative evaluation
4. Understand the difference between internal and external evaluation

PROGRAM EVALUATION VIGNETTE

An urban school district receives a three-year grant to implement an after-school program to improve student academic achievement. As staff start to implement the program, the district administrator realizes that an evaluation of their program is required. The district administrator also realizes that such work requires expertise of someone from outside the district, and the superintendent, with permission from the school board, hires an external evaluator from a local college. After reviewing the grant, the evaluator conducts an initial review of the program's curriculum and activities. Next the evaluator develops an evaluation plan and presents it at the next school board meeting. The evaluation plan overviews the objectives that the evaluator has developed and the tools that he will use to collect the data. As part of the data collection process the evaluator discusses how the plan will provide two different types of feedback. Formative evaluation will be used to address issues as the program is happening. For example, one question might be: *Are all the stakeholders aware of the program and its offerings?* Summative evaluation will be used to answer the overall evaluation question: *Did students in the after-school program have a significant increase in their academic achievement over those students who did not participate?*

The board approves the plan and the evaluator spends the following month collecting data for the formative and summative portions of the project.

At the next board meeting the evaluator presents some of the formative evaluation and reports that there is a need to increase communication with parents. He suggests that the program increase the number of fliers that are sent home, update the school website, and work more collaboratively with the parent council. In addition, he notes that there is a wide variation in parent education levels within the district and that a large number of parents speak Spanish as their native language. The evaluator recommends that phone calls be made to parents and that all materials be translated into Spanish.

At the end of project year one, summative findings are presented in a final report. The report shows that lack of parent communication is still a problem, and there is little difference in the scores on the standardized measures used to gauge academic achievement of those students who participated in the program and those comparable students who did not participate.

Based on the evaluation report, district officials decide to make modifications to the program for the upcoming year. A parent center, which was not part of the original plan, is added, in the belief this will help increase parent involvement. In addition, the administration decides to cut back on the number of extracurricular activities the after-school program is offering and to focus more on tutoring and academic interventions, hoping that this will increase academic achievement in year two.

WHAT IS PROGRAM EVALUATION?

A common definition used to separate program evaluation from research is that program evaluation is conducted for decision-making purposes, whereas research is intended to build our general understanding and knowledge of a particular topic and to inform practice. In general, program evaluation examines programs to determine their worth and to make recommendations for programmatic refinement and success. Although such a broad definition makes it difficult for those who have not been involved in program evaluation to get a better understanding, it is hoped that the above vignette highlighted some of the activities unique to program evaluation. Let's look a little more closely at some of those activities as we continue this comparison between program evaluation and research.

What Is a Program?

One distinguishing characteristic of program evaluation is that it examines *programs*. A program is a set of specific activities designed for an intended purpose, with quantifiable goals and objectives. Although a research study could certainly examine a particular program, most research tends to be interested in either generalizing findings back to a wider audience (that is, quantitative research) or discussing how the study's findings relate back to the literature (that is, qualitative research). Most research studies, especially those that are quantitative, are not interested in knowing how just one after-school program functioned in one school building or district. However, for those conducting program evaluations this is seen as precisely the purpose.

Programs come in many different shapes and sizes, and therefore so do the evaluations that are conducted. Educational programs can take place anytime during the school day or after. For example, programs can include a morning breakfast/nutrition program, a high-school science program, an after-school program, or even a weekend program. Educational programs do not necessarily have to occur on school grounds. An evaluator may conduct an evaluation of a community group's educational program or a program at the local YMCA or Boys and Girls Club.

Accessing the Setting and Participants

Another characteristic that sets program evaluation apart from research is the difference in how the program evaluator and the researcher gain access to the project and program site. As described in the vignette, the program evaluator was hired by the school district to conduct the evaluation of their after-school program. In general, a program evaluator enters into a contractual agreement either directly or indirectly with the group whose program is being evaluated. This individual or group is often referred to as the *client.*

Because of this relationship between the program evaluator and the client, the scope of what the evaluator wishes to look at could also be restricted by the client. To have the client dictate what one will investigate for a research study would be very unusual. For example, because of the nature of qualitative research, a qualitative researcher who enters a school system to do a study on school safety may find a gang present in the school and choose to follow the experience of students as they try to leave the gang. If a program evaluation was conducted in the same school, the evaluator might be aware of the gang and students trying to get out of the gang, and this might strike the evaluator as an interesting phenomenon, but the evaluator would not pursue it unless the client perceived it as an important aspect of school safety or unless gang control fit into the original objectives of the program.

Collecting and Using Data

As demonstrated in the vignette, program evaluation often collects two different forms of evaluation data: *formative* and *summative*. A further discussion about formative and summative evaluation is presented later in this chapter; essentially, the purpose of formative data is to change or make better the very thing that is being studied (at the very moment in which it is being studied). This is typically not found in most applied research approaches. Rarely would the researcher have this reporting relationship, whereby formative findings are reported back to stakeholders or participants for the purposes of immediately changing the program.

Changing Practice

Although program evaluation uses the same methods as research to collect data, program evaluation is different from research in its overall purpose or intent, as well as the speed at which it changes practice. The overall purpose of applied research (for example, correlational, case study, experimental) is to expand our general understanding or knowledge about the topic and ultimately to inform practice. Although gathering empirical evidence that supports a new method or approach is certainly a main purpose of applied research, this doesn't necessarily mean that people will suddenly abandon what they have been doing for years and switch to the research-supported approach.

In the vignette, we can see that change occurred much more rapidly through the use of program evaluation. Based on the evaluation report, administrators, school board members, and project staff decided to reconfigure the structure of the after-school program and to provide parents with a center in the hope of increasing parent involvement. In addition, it was also decided that many of the extracurricular activities would be eliminated and that the new focus would be on the tutorial component of the program—hoping to see even more improvement in students' academic scores in the coming year.

For another example, consider applied research in the area of instructional methods in literacy. In the 1980s, the favored instructional approach was whole language; however, a decade of research began to support another approach: phonics. Despite the mounting evidence in favor of phonics, it took approximately a decade for practitioners to change their instruction. In the early 1990s, however, researchers began to examine the benefits of using both whole language and phonics in what is referred to as a *blended approach*. Again, despite substantial empirical evidence, it took another ten years for many practitioners to use both approaches in their classrooms. This is admittedly a simplified version of what occurred; the purpose here is to show the relationship between applied research and practice in the speed (or lack of speed) with which systems or settings that they evaluate implement changes, based on applied research.

Although there are certainly many program evaluations after which resulting changes do not occur swiftly (or at all), one difference between program evaluation and research is the increased emphasis that program evaluation places on the occurrence of such change. In fact, there are philosophies and approaches in program evaluation that so emphasize the use of evaluation findings that they believe if the evaluation report and recommendations are not used by program staff to make decisions and changes to the program, the entire evaluation was a complete waste of time, energy, and resources (Patton, 1997).

Reporting Findings and Recommendations

Another feature of program evaluation that separates it from research is the way in which program evaluation findings are presented. In conducting empirical research it is common practice for the researcher to write a study for publication—preferably in a high-level refereed journal. In program evaluation, as shown in the vignette, the findings are presented in what is commonly referred to as the evaluation report, not through publishing in a journal. In addition, the majority of evaluation reports are given directly to the group or client that has hired the evaluator to perform the work and are not made available to others.

Formative and Summative Evaluation

Both quantitative and qualitative data can be collected in program evaluation. Depending on the purpose of and the audience for the evaluation, an evaluator may choose to conduct an evaluation that is solely quantitative or solely qualitative or a mixed-methods approach. In addition to using quantitative and qualitative data, a program evaluator also has the option of providing summative and formative evaluation within a project (see Figure 1.1).

The choice of whether to conduct a summative or a formative evaluation is not exclusively dictated by whether the evaluator collects quantitative or qualitative data. Many people have the misperception that summative evaluation uses exclusively quantitative data and qualitative data is used for formative evaluation. This is not always the case. Whether evaluation feedback is formative or summative depends on what type of information it is and when it is provided to the client (see Figure 1.1).

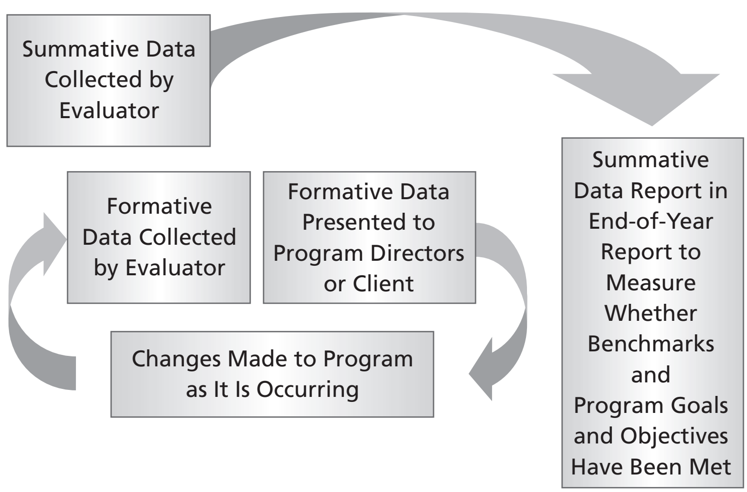

FIGURE 1.1. Formative and Summative Evaluation

Data for summative evaluation is collected for the purpose of measuring outcomes and how those outcomes relate to the overall judgment of the program and its success. As demonstrated in the vignette, summative findings are provided to the client at the end of the project or at the end of the project year or cycle. Typically, summative data includes student scores on standardized measures such as state assessments, intelligence tests, and content-area tests. Surveys and qualitative data gathered through interviews with stakeholders may also serve as summative data if the questions or items are designed to elicit participant responses that summarize their perceptions of outcomes or experiences.

For example, an interview question that asks participants to discuss any academic or behavioral changes that they have seen in students as a result of participating in the after-school program will gather summative information. This information would be reported in an end-of-project-year report. However, an interview question that asks stakeholders to discuss any improvements *that could be made* to the program to better assist students in reaching those intended outcomes will gather formative information.

Formative data is different from summative in that rather than being collected from participants at the end of the project to measure outcomes, formative data is collected and reported back to project staff as the program is taking place. Data gathered for formative evaluation must be reported back to the client in a timely manner. There is little value in formative evaluation when the evaluator does not report such findings to the client until the project is over. Formative evaluation feedback can be reported through the use of memos, presentations, or even phone calls. The important role of formative feedback is to identify and address the issues or serious problems in the project. Imagine if the evaluator in our vignette did not report back formative findings regarding parent communication. How many students might not have been able to participate in the after-school activities? One of the evaluator's tasks is to identify these program barriers, then inform program staff so that changes can occur. When programs are being implemented for the first time, formative feedback is especially important to developers and staff. Some programs require several years of intense formative feedback to get the kinks out before the program can become highly successful.

Formative feedback and the use of that information to change or improve the program is one factor that separates program evaluation from most types of applied research approaches. Classical experimental or quasi-experimental research approaches attempt to control for extraneous variables so that only the independent variable can affect the dependant variable. An important aspect of experimental research is a clear definition of the different treatments or level of independent variable. If the program itself is the treatment variable, then it must be designed before the study begins. An experimental researcher would consider it disastrous if formative feedback were given because the treatment was changed in the middle of the study. In contrast, program evaluators, while trying to keep the independent variables or treatment constant, realize that it is better to make modifications to the program—even if it "distorts" the lines of causality—than to deliver a substandard program consistently for the entire school year.

Training in Program Evaluation

Many students wonder *How do evaluators get involved in program evaluation?* and *Where do they receive their training?* These are both good questions. Although program evaluation today is certainly a much more recognized field, it is made up of both those who have formal training in program evaluation theory and practices as well as those who have been less formally trained. There is no specialized degree or certification required for people to call themselves evaluators. Today a number of colleges and universities offer course work in program evaluation as well as advanced degrees in this area. Although course work will vary by institution, most focuses on quantitative and qualitative methods, program evaluation theory, and ethics, as well as a practicum experience.

As in any field, program evaluators come from a wide range of backgrounds and experiences as well as different philosophical and methodological perspectives. Often faculty at colleges and universities serve as program evaluation consultants working with area school districts, agencies, non-profit and not-for-profit programs, and other institutions of higher education. There are also private evaluation consulting companies that hire program evaluators. Furthermore, public agencies at both the state and federal level also hire program evaluators for full-time positions to conduct internal evaluations in that setting, as well as to conduct single- and multisite evaluations.

The American Evaluation Association is an international organization devoted to improving evaluation practices and methods, increasing its use, promoting evaluation as a profession, and supporting evaluation to generate theory and knowledge. This organization has approximately four thousand members and representatives from fifty states and sixty countries. Each year the association hosts an annual conference in the United States that focuses on a theme, such as collaboration, methodology, or utilization (see http://www.eval.org/News/news.htm).The association also comprises special interest groups (SIGs) that specialize in certain areas or topics, such as teaching program evaluation or environmental evaluation.

INTERNAL AND EXTERNAL EVALUATORS

The proximity of an evaluator to what is being evaluated certainly plays a critical role in both the access to information, the collection of that information, and the reporting and use of that information to promote change. Just as the waiter at a restaurant has a very different perspective on the food and the management from that of the food critic who comes to dine and to write up a review for the local paper, an evaluator's perspective on and relationship to the setting or program must be considered. In the field of program evaluation, this perspective is often addressed by what are referred to as *internal* and *external evaluators*. An external evaluator is someone from outside the immediate setting who is hired to come in and evaluate the program. Since this person has no obligations, in theory he or she has no immediate biases for or against the program or any one of the stakeholder groups involved in the project. Most programs that receive federal, state, or foundation funding require an external evaluator to be present.

On the other hand, many companies, agencies, institutes of higher education, school districts, and other groups also employ internal evaluators. An internal evaluator is typically an employee of the company, agency, or group who is responsible for carrying out duties that pertain to evaluation. For example, many school districts now have a program evaluator on staff. This person is responsible for establishing and working with databases to maintain student academic and behavioral data and using data to assist staff and administration in improving practice. Internal evaluators at districts provide expertise in working with the state testing and accountability data, as well as monitoring programs that the school is currently implementing.

There are many strengths—and many barriers—to the use of both internal and external evaluators. As mentioned earlier, the main reason that many funding agencies require an external evaluator to be present is to increase the objectivity of the data collection. This may or may not be achieved; however, the external evaluator also inevitably encounters some barriers. External evaluators are often faced with the difficulty of establishing trust with the stakeholders involved in the program that they are evaluating. Even though the external evaluator is collecting data on the program and not specifically on the performance of program staff, this stakeholder group may not welcome the evaluator with open arms. Stakeholders may, and often do, see the evaluator as a threat—someone whose job it is to find "holes" in the program. They may see the evaluator's work as a direct threat to their livelihood. In some cases the stakeholders may feel that the external evaluator "really doesn't know us" or "doesn't know what we are all about." In some cases, they may feel that the evaluator doesn't know enough about the setting or the context of how things work in that setting to be able to gather in-depth data that would pertain to them and be meaningful for evaluation purposes. In many cases, stakeholders who are uncertain about this evaluator are likely to avoid the evaluator altogether, not returning phone calls to set up interviews or not returning surveys. It is a daunting and often difficult challenge for even the most seasoned of program evaluators to enter a foreign setting, establish trust with the various groups involved in the program, and provide meaningful data back to participants for programmatic improvements.

Internal evaluators typically do not have to deal with gaining the trust of stakeholders as external evaluators do. In addition, internal evaluators know the setting, how to access needed data, and the "language" that each group uses. In some cases both an internal and an external evaluator are retained. If an internal evaluator is already present in a school system or program, then an evaluation plan should encompass the work of both evaluators to optimize the breadth and depth of data collected and, ideally, to ensure the overall success of the program. In such situations, the internal evaluator would be responsible for collecting certain types of data to which the external evaluator would not have access. In turn, the external evaluator would collect additional data to ensure the authenticity and objectivity of the evaluation effort and its findings.

TYPES OR MODELS OF PROGRAM EVALUATION

Just as there are many types of applied research approaches, there are several different approaches that program evaluators can use. The most common approach to program evaluation is the *objective-based approach,* which uses objectives written by both the creators of the program and the evaluator. An evaluation objective is a written statement that depicts the overarching purpose of the evaluation and clearly states the type of information that will be collected. Often these objectives are further supported through the use of *benchmarks*. A benchmark is more detailed than an objective in that it specifically states what quantitative goals the participants in the program need to reach for the program to be successful. Box 1.1 presents an evaluation objective followed by a program benchmark.

BOX 1.1. Example of an Evaluation Objective and Benchmark

Evaluation Objective: To document middle school student changes in academic achievement, particularly in the area of reading and literacy skills.

Benchmark: Students in grades fifth through eighth will show a 10 percent gain on the ELA state assessment in year one and a 20 percent increase in students passing the ELA in program years two and three.

Evaluators will often start with the objectives for the evaluation and build evaluation data collection activities from those objectives. Evaluation objectives may guide either formative or summative data collection. Either way, quantitative or qualitative data, or both, are collected and findings are compared to the project's objectives. Objectives are certainly helpful in shaping the evaluation, but there is a risk that evaluators may become so focused on the objectives that they lose sight of other unanticipated outcomes or benefits to participants as a result of the program.

Although objectives assist in guiding an evaluation, there is another method—the *goal-free evaluation* approach—that doesn't prescribe using evaluation objectives. This approach is guided by the perspective that there are many findings and outcomes that do not fall within the strict confines of the goals and objectives established by both the project directors and the evaluator. Those who practice goal-free evaluation believe that the unforeseen outcomes may be more important than outcomes that the program developers employ. One difficulty in conducting a goal-free evaluation is that projects that receive funding are required to show specific outcomes based on objectives. If the outcomes are not included in the evaluation, the appropriate data may not be collected.

The *expertise-oriented evaluation* approach—one of the oldest and most frequently used methods of program evaluation—employs the evaluator to be a content expert and to serve more as judge than as evaluator

(Fitzpatrick, Sanders, & Worthen, 2004). Agencies granting accreditation to institutions, program, or services send program evaluators to these sites to conduct an expertise-oriented evaluation. In these situations, typically data are not collected by the evaluators but are presented to them by those participants being judged or seeking accreditation. With this approach, the evaluators judge the program or service based on an established set of criteria as well as their own expertise in the area. An example of this type of evaluation is the National Council for Accreditation of Teacher Education (NCATE). Colleges and universities who train teachers often seek national accreditation to demonstrate the quality of their programs.

The *participatory-oriented evaluation* approach uses a very different perspective on program evaluation than the other approaches described so far. Whereas those approaches focus on the program and examining different aspects of the program, the participatory-oriented evaluation approach is ultimately interested in those whom the program serves. Using this model, the evaluator will seek to involve program participants in the evaluation of the program. In some cases the participants will develop instruments, collect data, analyze data, and report findings.

HOW TO USE THIS BOOK

To provide some standardization, a framework was developed and applied to each case study in this book. Box 1.2 presents an overview of the framework sections and a brief explanation of each.

ADDITIONAL RESOURCES

As we can see, there are many different approaches to conducting an evaluation of a program. It should be noted that although the objective-based approach is not the sole approach for conducting an evaluation, because of federal and state funding and the focus on meeting goals and benchmarks in today's climate of accountability, it is, generally speaking, the most widely used approach. In addition, an objective-based evaluation will most likely be the first type of evaluation that a new evaluator just entering the trade will be exposed to and have to conduct. Therefore, most of the case studies presented in this book follow a more objective-based approach.

This section presents some additional resources and readings to assist those who are relatively new to program evaluation and to more

BOX 1.2 Overview of Framework to Guide Each Case Study

The Evaluator

In this section the evaluator or evaluators are introduced. The role of the evaluator is also discussed here, as well as the evaluator's background, education, and connection to the evaluation project as a whole.

The Program

Here the program that is being evaluated is described: its purpose, its implementation, and relevant stakeholders and participants. In addition, where possible, the goals and objectives of the program as well as the program's structure and design are overviewed.

The Evaluation Plan

Here the evaluator's evaluation plan is discussed in as much detail as possible. This includes but is not limited to the objectives that are driving the evaluation and the methods and tools that the evaluator is using or plans to use to conduct the evaluation.

Summary of Evaluation Activities and Findings

This section describes the data collection process of the evaluation and a summary or overview of any evaluation findings from the evaluation. In each of the cases, the evaluator(s) are usually presented with a dilemma or situation at the end of this section.

Final Thoughts

This section provides the reader with a conclusion: what really happened at the end of the evaluation, how the evaluator handled the dilemma, and the results of those actions for the evaluator and the project as a whole.

clearly delineate some of the activities overviewed and described in each case study.

The Evaluation Objective

In an objective-based evaluation approach, the evaluation objective is the cornerstone of conducting a rigorous and successful evaluation project. Evaluation objectives are written goals under which the evaluation data will be collected and reported. Box 1.3 presents a list of evaluation objectives used in evaluating the summer camp project.

BOX 1.3 Evaluation Objectives for a Summer Camp Project

Objective 1: To document stakeholder perceptions as to the purpose of the camp.

Objective 2: To document activities conducted during camp.

Objective 3: To document stakeholder perceptions of the lessons learned and the strengths and barriers of the camp.

Objective 4: To document student outcomes as a result of participating in the camp.

Objective 5: To document modifications made to programming based on the previous year's evaluation recommendations.

The typical evaluation has four or five main evaluation objectives. Specific data are collected to answer or address each evaluation objective. For many grant-funded projects evaluation objectives are already established and clearly defined in the grant. In such cases, an evaluator must work with the established objectives and begin to develop an evaluation matrix (see the following section). However, for projects with no pre-established evaluation objectives, the evaluator must play a significant role in their development.

Developing evaluation objectives in a collaborative setting can be a useful tool for an evaluator to use. To both build trust and gain buy-in from the different stakeholder groups (such as teachers, staff, administrators, parents), it is helpful to gather representatives from all parties for a discussion about the goals of the project and what outcomes or results they believe a program such as this would produce.

It should also be noted that evaluation objectives are not static; they can change over time. There may be objectives deemed important in the very beginning of a multiyear evaluation that are not emphasized at the end of the project. Typically, formative evaluation objectives (discussed shortly) are emphasized in the early stages of the evaluation timeline, and summative evaluation objectives (also discussed shortly) take on a more realistic and important role toward the end of the project.

No matter what objectives and timelines are being used, it is imperative that evaluation objectives are aligned to the project goals and activities. For example, let's say that the main focus of a summer enrichment program is literacy. As part of the program's activities, students or campers keep journals, work with local storytellers to author their own stories, and receive tutoring or interventions in literacy. From this experience project developers and staff hope that students will become more interested in reading and literacy as a whole and that this enthusiasm will eventually flow over to students' increasing their performance on some standardized reading measure they take at a later point. From this single program component, two evaluation objectives could potentially be developed, such as these two possibilities:

- *To document an increase in students' interest in and frequency of engaging in reading and other literacy-based activities.* Data for this evaluation objective could be collected through pre-post interviews with students documenting whether they believe their interest in and frequency of such practices has increased over time as a result of participating in the project. Supporting evidence could also be collected from parents, who may be observing their child reading more books at night, taking more books out of the library, talking about the book they are reading at dinner, and so on. An analysis of student campers' journals, the list of books they have completed, book reports, and so on could serve as additional evidence to support these claims.

The second objective could focus on more "hard" or end outcomes (such as test scores). A discussion of end outcomes is presented later in this section.

- *To document increases in student performance on a standardized literacy measure administered annually.* This objective would require the evaluator to obtain students' scores on the annual measure to determine whether there appears to be any relationship between students participating in the program and scores increasing on the assessment.

Evaluation objectives will vary somewhat depending on the program. However, there are some general categories that all objectives can fall under. Box 1.4 presents a description of those categories.

BOX 1.4. General Categories of Evaluation Objectives

Documenting Activities. Objectives such as these work toward documenting what the program "looks like" by describing what activities took place. Data for these types of objectives can be gathered through interviews, focus groups, or surveys (see "Tools for Collecting Data" later in this introduction), and through direct observations of program activities.

Documenting Program Implementation. These objectives focus on documenting processes associated with program start-up and basic program implementation. As part of this effort the evaluator would be interested in documenting strengths as well as barriers to program implementation. For example, one barrier might be that the evaluator discovers there isn't enough busing available for everyone who wants to attend field trips. Barriers that severely impact the quality of the programming (such as an instructor not using the correct curriculum) should be documented and fed back immediately to the project director so this problem can be corrected in a timely manner. Safety concerns are another barrier that requires immediate feedback. Again, evaluation in which information is presented to staff in a timely fashion is *formative*. Because of its timely nature, formative evaluation findings are often reported to program staff through the use of memorandum reports and presentations. These presentations can be done at the project's weekly or monthly meetings.

Documenting Outputs of Activities. These objectives focus on outputs or changes that occur as a result of some activity. These changes tend to be associated with what people *believe* or how they *perform* or *act.* For example, if program staff attended a seminar on working with at-risk students and as a result of engaging in this activity their beliefs about poverty changed or they changed some aspect of their instruction, this would qualify as a finding that would meet this objective. Data for these types of objectives can be gathered through interviews and surveys. Before using a survey to document these outputs, the evaluator should allow some time to pass between participating in the seminar and returning to the classroom. For an example of an output of activity objective, see the first of the previous set of two example objectives, "*To document an increase in students' interest in and frequency of engaging in reading and other literacy-based activities.*"

Documenting End Outcomes. These objectives focus on documenting changes in the participants themselves. In after-school and enrichment programs these end outcomes are often referred to as *hard outcomes*—that is, outcomes that are measured with a standardized measure or assessment; for example, changes in students' reading, math, or science scores on a standardized measure are considered to be end outcomes. A decrease in the number of violent incidences, increase in student attendance, and increase in student coursework grades could also be used to satisfy end outcome evaluation objectives.

Designing and Developing an Evaluation Matrix

One of the first activities to be conducted during the planning of the evaluation is the development of an evaluation matrix. The matrix serves as a blueprint to guide the evaluator and to ensure that all necessary data are collected. Table 1.1 presents an example of a matrix used to evaluate the summer camp project. Although each project will have its own unique evaluation objectives, the basic components essential to all evaluations are the same. Notice in the example matrix shown that the evaluation is being guided by five individual objectives. Notice also that the matrix contains the timeline for when the data will be collected and the methods and measurement tools the evaluator intends to use for data collection, and it specifies whether the data are summative (findings presented at the end of the project) or formative (findings presented as the project is occurring). The more detail the evaluator can present in the matrix, the easier it will be to conduct and carry out the overall evaluation. Most evaluators use some sort of matrix, even though it may not be spelled out as formally as the one in the table.

In addition to helping organize the evaluation, the evaluation matrix is also a wonderful tool in helping the evaluator build trust with the various stakeholder groups involved in the project. In doing so, evaluators may have early discussions with representatives from individual stakeholder groups (such as teachers, parents, staff) about the data collection process and the kinds of information that stakeholders perceive as important and useful. It is recommended that the evaluator incorporate the assistance and feedback from all stakeholders into the building of the evaluation matrix before data are collected. Keep in mind that on a multiyear project, the matrix and data collection activities are likely to change slightly as new objectives are added to the evaluation plan and old objectives that have been met and no longer need to be monitored are done away with.

Tools for Collecting Data

As specified in the evaluation matrix, the tools that the evaluator uses to collect data will vary depending on several factors, including the size of the stakeholder group, the educational or developmental level of the stakeholder group, and the evaluator's access to the stakeholder group. This section presents a few of the basic tools commonly used by evaluators and typical methodologies used for evaluations of after-school, summer, and enrichment-oriented programs.

TABLE 1.1. Evaluation Matrix for Summer Camp

Evaluation Objective	Stakeholder Group	Tools Used to Collect Data	When	Purpose
Evaluation Objective One: To document the depth and breadth of activities provided during follow-up session (2004–2005)	Faculty and project directors and campers	Interview	July	Summative
Evaluation Objective Two: To document student satisfaction with the follow-up activities	Students	Interview and observations	Now	Summative
	Parents	Post-survey	March–April	Summative
Evaluation Objective Three: To document faculty perceptions of the follow-up activities	Faculty and project director	Interviews	March–April	Summative
Evaluation Objective Four: To document parent perceptions of student outcomes from participating in camp and follow-up activities	Parents	Survey	March–April	Summative
Evaluation Objective Five: To document changes in student learning and abilities	Students	Word knowledge assessment	March 5 (post)	Summative

Data Sources

Surveys or self-report measures. The survey or self-report measure is perhaps the most common data collection tool used by program evaluators. One reason this tool is so popular is the overall ease with which surveys can be administered. Surveys are usually administered through a mail-out, mail-back procedure; however, in some cases they may be collected on-site, typically following an activity such as a workshop or an information session.

Surveys can be administered across multiple groups involved in a program. Keep in mind that in doing so, wording of items may need to be modified slightly for the different groups. The following is a list of stakeholders that the evaluator may want to consider surveying when conducting an evaluation of an after-school, enrichment-oriented, or summer program.

- Parents and guardians
- Project administrators
- Project staff
- Parents
- Community members, volunteers, senior citizens
- Students
- Presenters and service providers

Designing a Survey

When designing the survey it is important that the final form of the survey be piloted or field tested prior to being sent out, to ensure that there are no errors in the survey that would keep participants from being able to properly fill it out. In addition, it is important to be aware of possible language or reading ability barriers for those being surveyed. Pretesting the survey with a handful of those participants should give the evaluator an accurate idea of how the survey will perform when administered to the entire stakeholder group.

Exhibit 1.1 presents a survey designed to gather information from parents and guardians of the students participating in the project. The survey was specifically developed to address multiple evaluation objectives.

EXHIBIT 1.1. Parent or Guardian Perception Survey—Summer Camp

PLEASE RETURN by July 30

As part of the effort to evaluate the summer camp, the following survey has been designed to gather your perceptions regarding the activities associated with the camp. The information that you provide will assist in delivering important formative feedback to program coordinators and to the granting agency, as well as help in meeting the intended objectives and outcomes of the overall project. Your responses are confidential and will not be shared with anyone in any way that identifies you as an individual. Only aggregated data will be presented in the final evaluation report. Your participation in this survey process is completely voluntary and will not impact your child's future attendance in the program. Your time and cooperation are greatly appreciated. If you have any questions regarding this survey or the overall process please contact Dr. Dean T. Spaulding, Assistant Professor, Department of Educational Psychology, College of Saint Rose, Albany, NY 12203, (518) 454-2865 or e-mail at Spauldid@strose.edu

Please have the parent or guardian who had the most involvement with camp complete this survey.

Perceptions of Recruitment

The following items seek to gather your perceptions regarding your overall beliefs about the recruitment process for summer camp. Please read each item carefully and use the scale below to show your level of agreement with each item.

1=Strongly Disagree 2=Disagree 3=Slightly Disagree 4=Slightly Agree 5=Agree 6=Strongly Agree

I was provided with camp information in a timely fashion.	1	2	3	4	5	6
The program brochure provided me with a way to get additional information prior to enrollment.	1	2	3	4	5	6
I found the enrollment process to be easy.	1	2	3	4	5	6

How did you hear about camp? ____

Perceptions of Orientation

The following items are designed to gather your perceptions about the orientation process for summer camp. Please read each item carefully and use the scale below to show your level of agreement with each item.

1=Strongly Disagree 2=Disagree 3=Slightly Disagree 4=Slightly Agree 5=Agree 6=Strongly Agree

I believe the check-in process at orientation was well organized.	1	2	3	4	5	6

I left orientation feeling confident that my child was in good hands.	1	2	3	4	5	6
I believe dinner at orientation allowed me to meet the counselors/teachers my child would be working with.	1	2	3	4	5	6
I think having dinner with my child at orientation allowed me to be included in the camp experience.	1	2	3	4	5	6
The information session at orientation provided me with a clear understanding of what my child would be doing at summer camp.	1	2	3	4	5	6
I was encouraged to participate in camp activities throughout the ten-day program.	1	2	3	4	5	6
I was provided contact numbers and information.	1	2	3	4	5	6
I was provided enough information so I could attend camp activities and field trips.	1	2	3	4	5	6
Food was appropriate for children.	1	2	3	4	5	6
I enjoyed the Hudson River Rambler performance.	1	2	3	4	5	6

If you went to the dorms with your child either on orientation night or visited at a later date, please answer the next three questions:

I left the dorms feeling my child was in a safe place.	1	2	3	4	5	6
I felt the house was clean.	1	2	3	4	5	6
I felt the room would be a comfortable place for my child.	1	2	3	4	5	6

Perceptions of Parent Involvement During Camp

If you participated in the following activities, indicate your participation with a ✓:

Date	Breakfast	AM Session	Lunch	PM Session	Dinner	Field Trip
Monday 7/5						
Tuesday 7/6						
Wednesday 7/7						
Thursday 7/8						
Friday 7/9						
Saturday 7/10						
Sunday 7/11						
Monday 7/12						
Tuesday 7/13						
Wednesday 7/14						
Thursday 7/15						

(Exhibit 1.1 continued)

If you did not participate in any or all of the activities on the front page, please circle your reason (circle all that apply):

A. I did not have transportation.

B. I had other child care needs.

C. I had work conflicts.

D. I thought I would have to pay to participate.

E. I was not interested.

F. I did not know I could participate.

G. Other ____

____

Reflections of Camp

From what you have heard or observed from your child, what did your child like about summer camp? (check all that apply)

___Food	___Dorm room	___Class time
___Field trips	___Counselors	___Teachers/Professors
___Night activities	___Speakers/Guest lecturer	___Other campers
___Campers' cameras	___Final presentations	___Working on the computers

Other (please explain): ____

From what you have heard or observed from your child, what didn't your child like about Summer Camp? (check all that apply)

___Food	___Dorm room	___Class time
___Field trips	___Counselors	___Teachers/Professors
___Night activities	___Speakers/Guest lecturer	___Other campers
___Campers' cameras	___Final presentations	___Working on the computers

Other (please explain): ____

1=Strongly Disagree 2=Disagree 3=Slightly Disagree 4=Slightly Agree 5=Agree 6=Strongly Agree

I believe my child wants to come back to camp.	1	2	3	4	5	6
My expectations of camp were met.	1	2	3	4	5	6
I believe my child's expectations of camp were met.	1	2	3	4	5	6

Perceptions of Impact on Academics/School

The following items are designed to gather your perceptions about the possible impact attending camp may make on your child's academics and school-related work in the upcoming school year. Please read each item carefully and use the scale below to indicate your level of agreement with each item.

1=Strongly Disagree 2=Disagree 3=Slightly Disagree 4=Slightly Agree 5=Agree 6=Strongly Agree

I believe that this camp experience will help my child in school.	1	2	3	4	5	6
My child has been continuing activities experienced at camp.	1	2	3	4	5	6
I have noticed improvement in the way my child interacts with other children.	1	2	3	4	5	6
I plan to attend the follow-up sessions with my child.	1	2	3	4	5	6
I would be willing to send my child to summer camp next year.	1	2	3	4	5	6
I would recommend summer camp to other parents.	1	2	3	4	5	6

Demographic Items (Optional)

About you (Please check all appropriate items):

School district ___________________________

Grade level (Fall 2003) _____________ Child's age _____________

Child's gender ____ male ____ female

Did your child attend camp last year? ____ yes ____no

Total number of members within the household? ____
Number of children ____ Number of adults ____

Which camp did your child participate in?
____Storytelling ____American history ____ I don't know

Which residence hall did your child live in?
____Fontebonne____ Charter____McGinn____ I don't know

PLEASE PROVIDE ANY ADDITIONAL COMMENTS:

____

____

____

Other Types of Scales for Collecting Data Through Surveys

A successful survey asks only needed information and is easy and quick to complete. A survey that is too general and appears to be asking questions that have little or nothing to do with the project will quickly be dismissed by those who are expected to fill it out. A survey should collect only data that are essential and needed for the evaluator to complete the evaluation of the project. In addition, the evaluator should know exactly which questions or items on the survey are aligned with which objectives. For example, an evaluator should know that items 4 through 14 will answer evaluation objective one and items 15 through 26 will address objective two, and so on. Planning in such detail will ensure that only essential information is collected. This information can be included in an evaluation matrix.

The following are a few common scales and approaches that can be used to solicit information from participants.

Likert scales. These scales are commonly used in surveys. Respondents are presented with complete statements (for example, "I found the program increased students' interest in reading") and use an agreement scale to indicate their beliefs, selecting the number that best represents how they feel. Here is an example of a Likert scale:

1=Strongly Disagree; 2=Disagree; 3=Slightly Disagree;
4=Slightly Agree; 5=Agree; 6=Strongly Agree

Checklist. This is an easy way to gather breadth of information from participants. A checklist is essentially a list of possible answers that respondents check off if applicable. Although constructing a checklist is not difficult, generating such breadth of items can sometimes pose difficulties, especially if the evaluator is not fully aware of all the possible answers that would be appropriate. Sometimes conducting a few initial interviews with members from stakeholder groups can help the evaluator expand the checklist to ensure that it gathers valid data. It is also advisable to include an "Other" category at the end of the checklist. This allows respondents to write a response that was not posted on the checklist. (See Exhibit 1.1 for examples of checklists.)

Open-ended or free response. These items ask an open-ended question and expect respondents to give detailed answers. Unlike the other methods just described, open-ended items allow the respondents to describe "how" and "what" in much more depth. However, in constructing a survey it is important not to overuse open-ended questions. Too many open-ended items on a survey can deter participants from filling it out. Open-ended questions should be used appropriately. Data derived from them should be linked directly to answering evaluation objectives, and the evaluator should avoid putting open-ended items at the end of the survey just to fill in any extra blank space.

Demographics. A demographics section can be placed at the beginning or end of a survey to gather personal information about the participants. This information can vary widely depending on the purpose of the project. The survey in Exhibit 1.1 has limited demographics; additional possibilities include gender, age, marital status, years employed in current position, educational level, and annual income.

One-to-One Interviews

Although many of us probably have some idea of how interviews are conducted, we may not realize that they are more than simply asking questions of someone and writing down his or her responses. To have a successful interview requires proper advance planning. The evaluator needs to establish the time and location and develop a protocol or list of questions. These lists are often called *interview protocols*. Typically, an interview protocol contains no more than six to eight open-ended questions. Interviewing with such a list should take about an hour, depending on the project and the level of detail that is needed. As with the other tools, questions from the interview protocol must also be linked to specific evaluation objectives.

Aside from developing six to eight broad questions, the evaluator may also want to develop subquestions or *probes*. Probes help to ensure that specific information within the larger context of the questioning process are being addressed. One of the benefits of using an interview protocol in conducting multiple interviews is that the protocol helps to standardize the process, so everyone is asked the exact same questions, word for word.

Exhibit 1.2 presents an example of an interview protocol that was used for the evaluation of the summer camp. This protocol was used to interview instructors from the summer camp. Questions 3 and 7 provide examples of a subquestion or probe.

Another method of collecting data from stakeholders, the *focus group*, is very similar to one-to-one interviews. To conduct a focus group, the evaluator first develops a protocol—a series of open-ended questions; however, instead of asking them of an individual, the questions are posed to a group of stakeholders for discussion. The advantage of this technique is that often the conversations will get much deeper because of the different perspectives of the assembled individuals.

EXHIBIT 1.2 Interview Protocol

1. What was the purpose of the follow-up sessions?

2a. What was the overall process for developing the follow-up sessions?

2b. How does that extend and support the curriculum delivered at the summer camp?

3. Describe the activities used in the follow-up session. Which of these did you find the campers were most and least engaged in?

4. What do you see as the main learning objectives of the activities?

5. Overall, have the learning objectives been met? If so, how?

6. What changes would you make to the curriculum for next year's follow-up?

7a. What changes do you see, if any, in these students in the time you have been working with them?
- As a group?
- On an individual student basis?

7b. What other possible changes in students' performance could you expect to see as a result of students participating in this experience?

8. What do you see as the Saturday follow-up's strengths?

9. What do you see as barriers?

10a. Has your experience in developing and implementing curriculum for camp and the follow-up sessions changed how you think about or develop curriculum for your college classes?

10b. Has it changed how you instruct others to teach this population?

11. What are some of the lessons you have learned from this experience?

When conducting a focus group, it is important that the evaluator set ground rules beforehand to ensure that all participants respect each other, even if their views on the situation are very different. At least two evaluators should be present when conducting a focus group: one to ask the questions and the other to take notes.

A video or an audio recording device can be used during both interviews and focus groups. This will help to ensure the accuracy of the data being collected by allowing the evaluator to add further detail and quotes that might not otherwise have been recorded. If the evaluator is planning on using such a device, it is important that those being interviewed are informed and agree, both off and on tape.

Alternative Forms of Data

In addition to using surveys and interview protocols, evaluators are always seeking creative ways to collect different kinds of data. Often, when working with school-age children, evaluators will have the students keep journals about their experiences with the project. When considering using journals as a source of data collection, it is important—especially with students in middle school grades—to provide some sort of structure for their journal entries. One way to do this is to provide daily or weekly themes or even questions that students must respond to by writing in their journals. In addition, the evaluator should make it quite clear that the student journals are going to be collected and read as part of the evaluation.

Photography is another excellent method for collecting data. An evaluator who wishes to use photography as an alternative data collection method has several options. First, the evaluator can either choose to be the photographer and photograph students engaging in activities or allow the students to be the photographers. During the summer camp program, campers were each given a disposable camera and asked to photograph things that they liked or didn't like about camp. Over the course of the next ten days campers took lots of pictures during field trips, class time, and free time. Later the photographs were developed and students were interviewed, using their photographs as prompts to further the conversation.

Writing the Evaluation Report

There is no one way to construct an evaluation report, but there are some general guidelines. Typically, summative evaluation reports are written and presented at the end of each project year. In some cases, a midyear project status report is required. As the evaluator you should determine whether such a midyear project is needed and plan accordingly.

The following are the basic sections of an evaluation report:

Cover Page. This should contain the title of the project, the evaluator's name and credentials, the client or name of the organization who commissioned the report, and the date or time of year that the report is being submitted (for example, Summer 2005).

Executive Summary. For short reports, an executive summary is not necessary. Typically, an executive summary runs one or two pages and provides a short purpose and methodology for the report, the essential main findings, a conclusion, and recommendations, if appropriate. Often administrators use the executive summary as a stand-alone document to highlight key findings at meetings, media events, and the like.

Introduction. A two- to three-paragraph introduction is a good way to set the stage for the evaluation report and how the project came to be. In addition, the introduction should contain the overall *purpose* of the evaluation, the *name of the client* or organization for which the report is being written, and both the *project goals* and the *evaluation objectives.*

Methods. In this section the evaluator presents an overview of the different types of tools that were developed, when they were administered, what kinds of data were collected, and the sources for the data.

Body of the Report. The body of the report contains the analyzed data and findings from the evaluation. It is best if you start off each new objective on a new page. First, the evaluation objective should be restated, followed by another short description of what tools were used, what kinds of data were collected, and from whom. Following this, the evaluator will want to report the summarized data in a table or a figure or in bulleted form. Based on this information, the evaluator will then generate an evaluation finding or findings. These evaluation findings are generally an overall theme or summary of the data that are being presented. Additional data that supports the main data and findings can be presented in bullet form under the main table (see Exhibit 1.3).

EXHIBIT 1.3. Overview of Evaluation Objective and Finding

Objective 3: To document stakeholder perceptions of the lessons learned and the strengths of and barriers to the weekend follow-up sessions.

The purpose of this objective was to document stakeholder perceptions of both lessons learned from the experience and the strengths of and barriers to the weekend follow-up sessions. To meet this objective, qualitative data were gathered via either one-to-one interviews or focus groups. Parent perceptions data were provided through open-ended questions on the survey.

Finding: All stakeholders reported that maintaining friendships and becoming motivated to learn and build skills were the strengths of participating in the experience; lack of full participation and inconsistent attendance at follow-up sessions were noted by some to be barriers. Table 3.1 presents these findings by stakeholder category.

(Exhibit 1.3 continued)

Table 3.1. Stakeholder Perceptions of Strengths of and Barriers to Follow-up Sessions

Stakeholder	Strengths	Barriers	Suggestions
Program Director	• Continued friendships, making new friends • Exposure to students from different schools and background • Students continuing to learn and refine skills learned from previous lessons.	• Tutoring sessions occurring at same time • Families moving • Lack of contact information • Transportation • Conflicts with other school or family obligations	See if tutoring could come before or after; better integration
Camp Instructors	• Continued friendships, making new friends • Exposure to students from different schools and background • Students continuing to learn and refine skills learned from previous lessons	• Only 50% of students attending • Month of time between each follow-up session too long • Difficult to keep students on target with learning between sessions • Inconsistency with student attendance, difficult to provide continuity	Linking students together via Internet or Blackboard
Campers	• Able to see friends and stay in touch • Able to learn more about history and storytelling • Able to improve computer skills	• Sessions too short, need to be longer • Not all students attended	• Lengthen sessions • Mandatory attendance
Parents	• Continue learning and growing • Friends	Family or school obligations on same day; sessions were known about in advance so planning could occur	Overall, Saturday mornings a good time

(Exhibit 1.3 continued)

Finding: Camp instructors noted that this experience has benefited their own pedagogy and teaching at the college level, as follows:

- Examination of qualitative data revealed that camp instructors noted several areas in which their work serving as instructors for camp has benefited or changed how they think about or deliver instruction at the college level. More specifically, instructors noted that because of this experience they have tried to do more with interactive activities in their college classroom and have seen how effective such practices are when teaching an adult population.
- Camp instructors reported that this experience has also changed how they instruct others to work with urban at-risk youth. More specifically, instructors have gained this insight: it is important to stress to preservice teachers that when instructing students from these backgrounds they should allow for extra time to start an activity, as it takes these students a little more time to get into the activity.

Finding: Stakeholders noted several areas in which changes could be made to next year's programming in relation to the follow-up sessions:

- To address the issue of low attendance at follow-ups. During the initial greeting of parents at the summer camp, the follow-up sessions will be stressed, as well as their function to support and extend the work and learning that has occurred at summer camp. Parents will be reminded of these sessions at the closing of camp, and perhaps via a notice sent out at the beginning of the school year.
- Offer incentives for students to attend the follow-up sessions. Stakeholders believed that offering some type of incentive to students for completing the follow-up sessions would greatly help to increase the low attendance and to decrease inconsistencies in attendance that occurred with this year's sessions.
- Lengthen the time for sessions. Another area to be addressed is the time constraint with the current three-hour sessions. Stakeholders noted that combining two months of sessions would allow for a half- or quarter-day field trip to a museum or other appropriate educational venue.
- Increase parent involvement. Stakeholders also noted the need for more parent involvement in the follow-up sessions; they believed that field trips could be used as a way to get more parents involved.

Dissemination and Use of Evaluation Findings

It is the role and responsibility of the evaluator to deliver the evaluation report on time. The report should be delivered to the client or agency that has directly commissioned the work to be conducted. In the case of summer-enrichment programs, the client will most likely be an administrator or a project director (or both). In most cases it is the responsibility of the administrator or project director to submit the final evaluation report to the funding agency. Even if an evaluator has established trust and a positive relationship with a particular stakeholder group (such as parents), he or she cannot give the evaluation report to the group without the expressed permission of the client. Once the client has reviewed and made comments to the evaluator, the client will disseminate the report to whichever groups he or she feels should receive it. In some cases the client may wish to have the evaluator present the key findings from the executive summary at an upcoming project meeting and field any questions that stakeholders may have.

The appropriate use of evaluation findings and recommendations are key to a successful evaluation project. Ideally, throughout the process the evaluator has established a professional degree of trust among the stakeholders he or she has been working with. One of the silent roles of the evaluator is to present evaluation findings and recommendations to the client in such a way as to make change occur. The role of the evaluator does not stop with the delivery of the report and recommendations. The evaluator should work with the client to address these issues, and to continue to gather and feed data back to the client until they are resolved.

One way an evaluator can monitor progress toward meeting the recommendations for the project is to build this activity into an evaluation objective. As part of the evaluation of the camp, the evaluation team did just that: they built in a specific objective that focused on the project staff's ability to address limitations or concerns within the project. At the end of the camp all areas of concern had been successfully addressed. Exhibit 1.4 presents this objective.

EXHIBIT 1.4. Example of Evaluation Objective Focused on Program Modifications

Objective 5: To document modifications made to programming based on year one program evaluation recommendations

The purpose of this objective was to document any programmatic changes that were conducted in year two that were based on program evaluation recommendations from year one. To complete this objective, a review of the year one follow-up report was conducted. In addition, qualitative data were gathered from stakeholders and data across the entire report were analyzed to determine

(Exhibit 1.4 continued)

whether program refinements had been made and whether they were successful. Finding: In 2004-2005, all recommendations made from year one were addressed and intended outcomes were achieved.

Table 5.1. Status of Prior Recommendations Made to the Summer Camp Follow-Up Sessions

2003-2004 Recommendations	2004-2005 Changes	Results	Status
Increase interest and awareness of follow-up sessions during summer camp 2004.	An emphasis was made by prior campers and staff to increase awareness of follow-up sessions.	There was a 50% increase in total number of campers attending follow-up sessions.	Achieved
Decrease number of sessions, increase length to include trips.	Number of sessions was shortened from six to five total.	Campers attended a full day trip to Boston.	Achieved
Provide field trip opportunities.	• Five Rivers—snowshoeing • Albany—Underground Railroad tour • Boston	Campers realized that learning can take place outside of a classroom environment.	Achieved
Provide an incentive for completing follow-up activities.	Culminating activity was a trip to Boston's aquarium, IMAX, and planetarium.	A total of thirty campers attended the culminating activity.	Achieved

KEY CONCEPTS

Client

Formative evaluation

Summative evaluation

Benchmarks

Program

Internal evaluators

External evaluators

Objective-based approach

Goal-free approach

Expertise-oriented approach

Participant-oriented approach

SUGGESTED READING

Cousins, J. B., & Earl, L. M. (1992). The case for participator evaluation. *Education Evaluation and Policy Analysis*, *14*(4), 397–418.

MacNeil, C. (2002). Evaluator as steward of citizen deliberation. *American Journal of Evaluation*, *23*(1), 45–54.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.