



FEDERATION FRN FOUNDATION CUBE NATIONAL AFFILIATE SCHOOL ATTORNEYS TLN CAUCUSES

Program Evaluation

[About EDLO](#) → [EDLO Site Map](#) → [Contact Us](#) → [Help](#) →

[Home](#) // [Evaluation](#)

"A" is for Assessment: A Primer on Program Evaluation

"A" is for Assessment: A Primer on Program Evaluation

Introduction

The Nellie Mae Education Foundation promotes access to quality education for historically underrepresented low-income students and adult learners in New England through strategic grantmaking, research, advocacy, and technical assistance, including program evaluation.

We believe that evaluation is a critical part of program design and is necessary for ongoing program improvement. Evaluation requires collecting reliable, current and compelling information to empower stakeholders to make better decisions about programs and organizational practices that directly affect students. A good evaluation is an effective way of gathering information that strengthens programs, identifies problems, and assesses the extent of change over time. A sound evaluation that prompts program improvement is also a positive sign to funders and other stakeholders, and can help to sustain their commitment to your program.

We hope this guide helps practitioners and others strengthen programs designed to increase academic achievement, ultimately broadening access to higher education for youth and adults. This guide was developed by Policy Studies Associates, Inc. with editorial assistance from the Foundation's staff.

Why Evaluate?

Evaluation is the key to program improvement because it shows which areas of a program are strong and which need to be strengthened. Good evaluation is a tool program managers and other stakeholders use to:

- Measure the effectiveness of activities
- Document changes in programs and participants
- Monitor whether programs are on target to reach goals
- Inform program staff of effective strategies and areas needing improvement
- Modify and improve activities to reach desired outcomes
- Be accountable to stakeholders including the community, funders, schools, students, and parents.

Theory of Change in Evaluation

Theory-based evaluation is an analytic process that helps determine what is working in a given program or project and what needs to change.

Theories of change are conceptual maps that show how and why program activities will achieve short-term, interim, and long-term outcomes. The underlying assumptions that promote, support, and sustain a program often seem self-evident to program planners. Consequently, they spend too little time clarifying those assumptions for implementers and participants. Explicit theories of change provoke continuous reflection and shared ownership of the work to be accomplished. Even the most experienced program planners sometimes make the mistake of thinking an innovative design will accomplish goals without checking the linkages among assumptions and plans.

Developing a theory of change is a team effort. The collective knowledge and experience of program staff, stakeholders, and participants contribute to formulating a clear, precise statement about how and why a program will work. Using a theory-based approach, program collaborators state what they are doing and why by working backwards from the outcomes

they seek to the interventions they plan, and forward from interventions to desired outcomes.

When **defining a theory of change**, program planners usually begin by deciding expected outcomes, aligning outcomes with goals, deciding on the best indicators to evaluate progress toward desired outcomes, and developing specific measures for evaluating results. The end product is a statement of the expected change that specifies how implementation, resources, and evaluation translate into desired outcomes.

Continuously **evaluating a theory of change** encourages program planners to keep an eye on their goals. Statements about how and why a program will work must be established using the knowledge of program staff, stakeholders, and participants. This statement represents the theory underlying the program plan and shows planners how resources and activities translate to desired improvements and outcomes. It also becomes a framework for program implementation and evaluation.

Developing Indicators

Once program expectations and theories of change have been established, program staff can **identify appropriate indicators** that will show whether or not expected outcomes are being achieved. The following questions may be asked of a program:

- What outcomes can be reasonably expected given existing resources?
- What contextual factors (beyond the control of program activities) might affect outcomes?
- What information is needed to improve the program?
- What is the most accurate measure of outcomes?
- What resources and measures are available for data collection?
- How can proposed data collection methods support program implementation?

Common Evaluation Language

Before embarking on evaluation, it is important to understand key terms to ensure that everyone involved applies definitions and standards consistently when collecting data.

Common evaluation concepts are discussed briefly below.

Conducting a **needs assessment** is often the first step in developing a program. This can be done through surveys, focus group discussions, observations, and interviews to identify problems and concerns in a community or segment of the population. A typical needs assessment also describes resources that are available to solve the problem, and helps stakeholders make informed decisions about program goals and what is required to accomplish them.

After defining needs, stakeholders design a program to address those needs and clarify goals. Goals are the desired outcomes of a program—the end toward which a program or set of activities is directed.

Types of Evaluation Strategies

Evaluators typically use at least three types of evaluation strategies for monitoring program implementation: monitoring, implementation, and outcome evaluations. A fourth type, impact evaluation, is typically undertaken for in-depth and long-term programs.

1. Monitoring evaluation is usually conducted internally by program staff to strengthen program design and management by measuring how accurately a plan has been executed; how well program resources have been invested; and whether adjustments in investments, resource allocations, or activities are necessary. Monitoring evaluations are conducted once or twice a

year and answer the question: Is the intervention implemented consistently with the program's design and goals? If not, why?

2. Implementation evaluation (sometimes called process evaluation) examines the implementation process, e.g., whether activities have served the expected number of people and provided an appropriate amount and quality of services. It explores how activities, programs or groups of programs have been conducted. This type of evaluation is often ongoing, enabling program staff to see the results of activities against process indicators and determine if the program is progressing towards its indirect goals.

3. Outcome evaluation focuses on the program's impact on the participants. Ideally, outcome evaluations occur annually to answer the question: Is the intervention achieving its intended effect? An outcome evaluation might look at whether completing a college preparatory program encourages more students to enroll in four-year universities. Outcome evaluations are often conducted by an outside evaluator and result in lessons learned that can be generalized beyond program participants.

4. Impact evaluation answers the question: What outcomes did participants achieve in this program over and above what would have occurred had the program not intervened? This type of evaluation usually uses formal scientific methods including control groups, sampling strategies, and statistical procedures to clarify the role of various program features and generalize the value of the program beyond its immediate impact.

Other key terms relevant to program evaluations include:

- **Outcome:** a particular goal the program hopes to achieve through strategies and activities. An outcome for a program that upgrades students' mathematics skills to prepare them to take advanced mathematics might be that the students take advanced mathematics classes in high school or college.
- **Indicator:** a specific, measurable display of an outcome. In the

example given above, an indicator could be students' math grades in advanced mathematics or scores on advanced placement tests.

- **Measure:** the actual tool used to collect information on the indicator. Enrollment records, student report cards, and test scores can be used to collect data about the continuing effects of a mathematics preparation program.
- **Benchmark:** a target for an indicator, usually based on best performance. The benchmark could be that half the students participating in an afterschool enrichment program received a "B" or better in English on their report cards. It may be a progress point toward a larger goal—a mid-point rather than an end point.

Measuring Outcomes

Outcomes are the valued result of programs. There are many ways to measure them and different kinds of information to gather about them. Sometimes, however, it takes a long time to achieve ambitious outcomes. In those cases, it may be important to define and measure both early and interim outcomes, while giving long-term outcomes/results more time to root. Measuring incremental changes can show if a program is on target and whether the changes occurring are the desired ones. Interim outcomes also indicate the validity of a program's theory of change and its relationship to goals, activities, and benchmarks. The most useful evaluations assess interim outcomes at least once a year—reporting their results internally for program improvement.

The most useful evaluations gather both quantitative and qualitative data. Both contribute to understanding program implementation and its effects. Evaluators use different strategies for

collecting, analyzing, and reporting different types of data.

Quantitative data is information that can be counted and measured, i.e., test scores, number of people who participate in a program, enrollment levels, and graduation rates.

Qualitative data are descriptive and interpretive. They include words, pictures, graphs, and designs that describe how the program works and affects participants. This type of information is typically gathered through interviews, focus groups, observations, and surveys with open-ended responses.

Evaluation Strategies

Evaluating a program need not be overwhelming. The place to start is with a clear understanding of the evaluation's purpose and the information needed to continually improve the program. Three simple steps help evaluators design evaluation strategies:

- Decide whether to use an internal evaluator or an external evaluator.
- Clarify the purposes of the evaluation.
- Involve grantees and other stakeholders in the design.

Internal Evaluator vs. External Evaluator

Depending on what resources are available and the type of evaluation being planned, you may choose to use an internal evaluator or an external evaluator. An **internal evaluator** is a member of the program staff or an individual hired expressly to oversee the evaluation process. This person may be assigned to work exclusively on evaluation or may have other program responsibilities as well. An advantage to this type of evaluator is that she or he may

be more familiar with the staff and program activities, have greater access to program resources, and wider opportunities for ongoing evaluation activities.

An **external evaluator** is a person or organization contracted from outside the program.

External evaluators may have greater access to outside resources. They may also have more expertise and experience in evaluation.

The kind of evaluation expertise best suited to a particular program will vary depending on the type of evaluation to be conducted, internal expertise, and resources. For instance, a program that relies primarily on surveys and participant interviews for evaluation data might prefer to hire an evaluator with experience designing and analyzing surveys and interview protocols. If the evaluation will focus on student progress as measured by test scores, an evaluator with strong quantitative skills and knowledge about student assessment will be most helpful.

In general, consider whether the candidate or evaluation group has the following qualities:

- Expertise in evaluation theory and methods
- Knowledge about the practice, program, or policy being studied
- Expertise in developing and implementing evaluations
- Ability to lead and manage a team of professionals from different fields
- Substantive record of conducting similar studies and analyses
- Confirmed reputation for being a collaborator and working in a timely manner
- Skill in helping to identify goals and theories of change
- Ability to develop data products that meet stakeholders' information needs and can be used to communicate with various audiences
- Flexibility to accommodate changing program and evaluation needs.

Clarifying Evaluation Purpose

Before beginning an evaluation, there should be consensus about why it is occurring, what

will be evaluated, and what all stakeholders and participants expect to learn from it.

Evaluation proposals, designs, and memoranda of understanding are useful devices for communicating and clarifying these issues.

Evaluation plans should specify:

- Who will conduct the evaluation
- Questions to be answered
- Timeline for evaluation activities
- Added value to be provided by the evaluation
- The budget.

Periodically reviewing an evaluation's rationale and goals keeps the process relevant.

Involving Stakeholders in the Evaluation Design

Stakeholders typically have an investment in program evaluation, making communication an essential ingredient in the evaluation plan. Inviting stakeholders, including grantee staff members, community advisors, and key foundation staff to contribute to the design of an evaluation insures it will address everyone's questions. The same questions and data collection strategy can satisfy varied purposes and audiences. Participation by stakeholders strengthens data collection, self-assessment and information management -- all of which contributes to a stable, sustainable program.

Planning the Evaluation Strategy

A number of considerations go into deciding on the best evaluation strategy. Planners need to know if there are adequate financial and human resources available to accomplish the task. They should also consider how they will use the information to be collected and determine if the burden placed on participants or other stakeholders is appropriate. A carefully planned evaluation strategy should inform program staff and funders, as well as promote program

improvement without overwhelming either participants or program staff.

Strategies for making evaluation an effective tool for program improvement include:

- **Begin the process early.** Integrating evaluation into program development facilitates documentation of baseline characteristics of participants and information about program activities and management. Baseline information serves as a reference point for changes that occur as a program develops over time.
- **Collect data annually.** Changes in learning and behavior occur slowly. It is usually not necessary to collect data more than once a year or at the completion of a program segment, whichever comes first, to draw an accurate picture of program effects. Sampling strategies can reduce the data collection burden and cost without changing what can be learned.
- **Use quantitative and qualitative data collection and evaluation strategies.** Quantitative data, i.e., report cards, test scores, administrative databases, and qualitative data (from interviews, observations, focus groups) complement each other and provide useful information on different aspects of a program.
- **Communicate results.** Evaluation results should not surprise participants or stakeholders. Interim memos and reports convey lessons learned as the evaluation is implemented. Everyday language rather than jargon makes information easy to understand and minimizes misunderstanding.

Data Collection Techniques

A program evaluation strategy usually involves a variety of data collection methods and sources. When designing your evaluation plan, consider whether the techniques you plan to use are appropriate for your program population. If participants are not fluent in English, for

example, a written survey will not be the best way of collecting useful information. A focus group might be a better method.

Clarity of language is essential in any data collection instrument. Make sure the instructions and questions asked are easily understood. It is helpful to pilot test a survey or interview on a subset of participants to determine if the questions asked generate accurate and complete information.

Target Population

The outcomes to be evaluated and the purpose of the evaluation will determine the target population. Evaluation planners should consider who will benefit from the program, then decide who or what to evaluate. Will all participants be asked to provide information about the program at some point? Will supportive stakeholders be invited to contribute to the evaluation? Will staff, participants, or some combination of both be interviewed or assessed, or will only a sample be involved? Will results from participants and non-participants be compared?

Other questions to consider include:

- How will the evaluation population be informed about the purpose, approach, and timeframe of the evaluation?
- Have data sources been informed about the evaluation (and its) schedule?
- Does the data collection timeline provide enough time to schedule interviews or a survey administration schedule?
- How will evaluation participants be compensated or thanked for their participation?
- If key informants speak a language other than English, what arrangements have been made for translators or translation of written materials?

Data Sources

Evaluating a program usually involves using a variety of data collection methods and sources. Data collection methods should be customized to the program and the population it serves. If a program is implemented for the first time, for example, it is proper for implementation concerns and indirect or early outcomes, rather than direct or long-term outcomes, to dominate start-up. If participants have low literacy levels or speak little or no English, conducting a focus group in participants' native language or asking a native-language speaker to translate will surely generate more accurate information than administering a written survey.

Data sources can be grouped into several categories: 1) documents and administrative records; 2) focus groups and interviews; 3) surveys; 4) observations; and 5) assessments of student knowledge. A combination of these sources usually provides the range of information that a program team and other stakeholders need.

Documents and Administrative Records

Grantees and program staffs are sometimes surprised to learn they do not have to collect all new data for an evaluation. Much of the information needed may be available in existing documents and administrative records. The following types of records can inform an evaluation.

Routine program documents often detail a program's processes and outcomes, indicating how a program changes over time. Relevant documents include mission statements, budgets, participant records, flyers, curriculum guides, schedules, staff guidelines, and annual reports. Examining documents and administrative records may also signal important questions for evaluators to ask in interviews or surveys.

Administrative records, e.g., lists of participants, activity descriptions, budgets, and

attendance records contain a wealth of quantitative information. They are especially useful in collecting information about program characteristics or events that evaluators cannot readily observe like the planning that takes place when the evaluator is not on site, or the effects of policy decisions on program changes. These records may come from multiple sources including the program and its agency collaborators or clients.

Examples of existing administrative records that may contain useful evaluation data include:

- School attendance reports
- Demographic records
- Community surveys
- School or district disciplinary records
- Participant report cards
- Curriculum materials
- State and national test scores
- Survey information collected by the community or state
- Extracurricular program reports.

Before beginning an evaluation, it is a good idea to decide what records you may need and allow time to negotiate access to them. Trying to obtain records at the last moment without the prior approval of administrators or clients can cause delays that will threaten the execution of even the best evaluation plan.

Focus Groups, Interviews, and Surveys

Focus groups, interviews, and surveys of people whose opinions may provide useful insights and information are excellent sources of information when evaluating programs. Good “people sources” include:

- Program participants

- Parents
- Program graduates
- Program staff
- Community agency members
- School staff and administrators.

A focus group is usually composed of 6 to 8 people with similar roles (e.g., teachers, students, parents) who meet to discuss topics posed by a facilitator. The group meets for one or two hours. One advantage to this method is that the group interaction may prompt people to think about things they might not have thought of on their own. A disadvantage is that some people may be uncomfortable expressing their opinions publicly, or may not want to contradict what someone else says, or share what they consider to be sensitive information. Giving individuals an opportunity to speak privately with the facilitator can offset this disadvantage.

Interviews can provide in-depth information about stakeholders' views of a program that is often difficult to obtain through other means of data collection. The following are some possible interview formats:

- There are no predetermined questions in an **informal conversational interview**. Topics come up during the course of conversation. This format insures that the questions are relevant to interviewees. It can generate a great deal of information, but it may also make getting a systematic view of a program difficult.
- An **interview guide** (or guided interview) outlines topics for discussion in advance. This method is more systematic, but still conversational. A drawback is that the interviewee may have valuable insights not covered in the guide.
- Questions are predetermined in a **standardized interview**. This method facilitates comparing responses, but limits the flexibility of the interview. Additionally, all of the questions may not be relevant to all interviewees.

- A **closed-field response interview** is the most structured interview method with pre-determined questions and answer choices. This method may seem impersonal, but it can be an efficient way to gather easily analyzed information. It may be an appropriate alternative to using a written survey of a low-literacy population.

It is a good idea to clarify what information you want by making a “protocol,” or list of questions, to guide and focus the conversation, increasing the likelihood that the data collection will yield valuable information.

Surveys are a direct way of measuring program outcomes. A survey can elicit data through open-ended questions, closed-ended questions, or a mix of the two. Closed-ended questions ask respondents to choose the most appropriate answer from a multiple-choice list. While closed-ended questions can be easier to analyze statistically, the pre-determined answer choices may not reveal subtleties in respondents’ opinions. Open-ended questions give participants more flexibility in sharing their thoughts about a program. Consequently, they make analysis more complex and take more time for respondents to complete.

Survey questions can assess perceptions, experiences, or current conditions. They can also assess client or participants’ changing views or behaviors. A popular survey technique is to ask participants to rate program quality or learners’ progress.

In general, survey questions should:

- Use simple language.
- Target one idea.
- Provide a range of possible answers in a closed-ended survey
- Be unbiased.
- Include clear instructions on how to answer questions and make sure questions are asked in a consistent manner.

Observations

Observation is another useful form of data collection, giving evaluators contextual information about a program and insights into its strengths and weaknesses. Observation can be a first step to designing relevant focus group questions, interview guides, and surveys. It also provides an opportunity to observe the program setting, activities, and interactions between staff and participants. A checklist of interactions and activities is a useful method of structuring observation.

Unobtrusive measures consist of systematic, but not obvious, observations of program activities and interviews with participants. They use no formal surveys or questionnaires. For example, they rely on periodic documentation of classroom activities; evidence of the need to reorder project materials or books; the thickness or number of pages completed in journals; or changes in the number of participants who use the library or log on to a web site. It is important to note for ethical reasons, however, that program staff and participants must be aware that an evaluation is in effect, even if they agree to not knowing specifically when data are being collected.

Evaluation Data

Once evaluators have settled on which data to use, they should decide how to collect it to insure that findings reflect the goals and objectives of the evaluation.

Pilot Testing Data Collection Instruments and Procedures

Data is easier to collect, more usable, and easier to analyze when evaluators plan and pilot test their instruments, sampling procedures, sources, and collection procedures. In any data

collection instrument, the clarity of the instructions and questions will affect the quality of the information gleaned. It is helpful to pilot test a survey or interview protocol on a subset of participants before using it on a large group to get feedback on whether the questions are clear and relevant. Evaluators should have confidence that data collection methods and tools will produce the needed information before using them in a formal evaluation.

Questions to consider:

1. Is it necessary to administer a survey to an entire classroom or can a smaller sample be assessed?
2. Are parents easily reached and willing to participate in a telephone interview or would an on-site interview give better results?
3. Are the procedures for administering instruments easy to follow?

Knowing the answers to these questions before data collection begins can save time and money.

Sampling Procedures and Comparison Groups

At this point, a decision should be made about sampling procedures. Systematically sampling a small number of participants or clients can reduce the data collection burden and expense associated with an evaluation. When considering which sampling procedure to use, keep in mind the following:

1. The sample size must be large enough.
2. The method for selecting the sample must be objective.
3. The population of interest for the data collection should be represented in the sample group.

Sometimes evaluators want to compare data on a program's results with data reported by a "comparison group" to see if there are systematic effects on participants. In order to draw valid conclusions, comparison groups should be similar to the population being studied, although obtaining records from groups that did not benefit from a program's services may prove difficult. Additionally, using a comparison group can clarify and validate findings, but

it adds complexity and expense to the evaluation process.

Obtaining Adequate Response Rates

Response rate is a critical factor in the validity of data collection. Information needs to be obtained from an adequate proportion of clients or participants. Fifty percent or more of an initial sample should respond to the data collection. The ideal is 80 percent, but it is unusual to achieve this level of return. Experts consider a higher response rate from a small sample more valid than a large number of returns from a non-representative population. Response rates lend credence to findings and should always be included when reporting evaluation findings.

Mail and telephone surveys typically yield lower response rates than face-to-face surveys, even when prizes or rewards are offered as incentives. One way to optimize response rates is to collect data face-to-face, perhaps when participants are gathered for meetings or conferences.

Finalizing Data Collection Instruments

Before beginning data collection, evaluators should re-examine three features of the data collection instruments:

1. Are the items aligned with the primary questions the evaluation is asking? Make sure every item on the instruments is pertinent to one of the evaluation questions.
2. Do the data collection instruments address all of the evaluation questions or is a key piece of information missing? If information is missing, now is the time to develop questions to address it. Conversely, there may be another data source that answers some of the evaluation questions. If so, select one to avoid redundancy.

3. The third feature to examine is appearance. If instruments look complicated, are hard to read or too lengthy, respondents will be put off and may not provide reliable information or may not complete them. Spacing, text size, correct grammar and spelling should all be checked. Investing in good printing and high-quality paper in an appealing color that can be duplicated easily can also make a difference in the quality of responses.

Firming Up Data Collection Procedures

Before beginning an evaluation, the following questions should be answered:

- Who will manage evaluation activities? Overseeing data collection and coordination consumes a great deal of time — staff will need to make sure evaluation instruments are copied and delivered to the appropriate location, as well as prepare a database to store and manage the information collected. A “point person” should be appointed to maintain regular communication between the grantee or program team and the evaluator.
- Who will administer the evaluation instruments and prepare the data for analysis? Data collectors need to be trained so that they are familiar with the key questions and purposes of the evaluation. Training data collectors also helps insure that evaluation results will be comparable regardless of who administers a survey or conducts an interview.
- When will data collection begin? Timelines should be established and data collection scheduled.
- What is the deadline for all data collection? All data must be collected before analysis can begin.
- When is the report of findings due?
- When will evaluation and program teams meet to oversee progress?
- Are there interim data reporting points or will follow-up information be collected from participants?
- Will data be collected longitudinally over several points in time?

Assuring Confidentiality

An evaluator has an ethical responsibility to assure confidentiality to anyone who provides information for an evaluation. Promising confidentiality also increases the likelihood that participants will refrain from giving what they think is the “right answer,” making the information collected more accurate and unbiased. The following suggestions can help meet an evaluator’s confidentiality commitments:

- Inform all respondents about confidentiality guidelines before beginning data collection.
- Make surveys anonymous or make clear how confidentiality will be ensured.
- When records differentiate individual respondents, use a neutral identifying code to eliminate the use of participants’ names.
- Do not use respondents’ names or other identifying information in any evaluation reports without permission.
- Keep all information secure.
- Periodically dispose of data no longer needed.

In addition to these precautions, evaluators may be requested to obtain permission from program participants before they begin to gather data from them. Student test scores, grades and attendance records cannot legally be collected without explicit permission from the students, or in the case of youth under 18, from their parents.

In some cases, evaluators may need to obtain permission from school officials or an “institutional review board,” known as an IRB. An IRB is a group within a school, university or other public institution that formally reviews, approves, and monitors research in accordance with federal, state and local regulations. Its purpose is to insure that researchers and evaluators protect the rights and welfare of individuals who participate in research studies.

Summarizing and Analyzing Data

Once data has been collected, the next step is to summarize and analyze the findings in a report to program officials and other stakeholders. Evaluations do not necessarily have to be statistically complex analyses to be effective. Simple approaches often work just as well. The following are descriptive, rather than statistical, methods of analyzing data:

- Organize qualitative data notes and other information by topic. Identify common themes about participants' experiences, success, and recommendations for program improvement.
- Consider the data in relation to the program environment. What are some possible explanations for the findings? What program factors might contribute to the outcomes suggested by the data? Do the results make sense?
- Reflect on the information in relation to the whole program. What occurred? How did findings contrast with expectations? How did behavior differ among groups or in different settings? Why? What program factors might contribute to the outcomes the data suggest? Do the results make sense?
- Focus on outcomes. Whenever possible, use tables or graphs to show the results of program participation.
- Explain findings and their implications by quoting program participants to illustrate the quantitative findings reported in a graph or table. Why did the results occur as they did?

The following are statistical approaches to using quantitative data:

- Convert findings into percentages or averages. For example, what percentage of students in the minority high achievement initiative went on to a four-year college? What percentage of participants in the adult literacy program received a high school equivalency diploma?
- Look at progress over time. What percentage of participants in an out-of-school program showed an increase in their literacy or mathematics grades?
- Compare participants with non-participants when possible. How did participants score on a statewide test compared to students who did not participate in the program; compared to other students in the district; in the state?

- When appropriate, describe the statistical significance of findings. “Statistical significance” is a concept that indicates what the likelihood is that the observed findings occurred by chance or would be the same if the data were gathered again using the same procedures. There are special procedures for determining statistical properties of data. Evaluators should consider when it is useful and appropriate to report them.

Do not over-generalize results. When interpreting outcome data, it is a mistake to assume the program is responsible for all of the positive changes. Many factors contribute to change and success in an education environment. It is important to acknowledge the possibility of other factors impinging on results.

Non-programmatic factors can influence data too. For example, did only the most involved participants respond to a survey? What happened to non-respondents? If an evaluation only tracks participants who succeeded in the program and stayed in touch, what can be learned about those who failed to complete the program? And what program changes could be made to insure a higher completion rate?

Finally, consider qualitative and quantitative results together to suggest what grantees can do to improve a program.

Using Evaluation Results

Effective evaluation assesses whether a program is successful in achieving its goals. It identifies effective practices and informs future planning, including strategies for program improvement. This information leads to program improvement in two ways:

1. Provides an analysis of program strengths and weaknesses to help program managers develop strategies for improvement
2. Can be used as an advocacy tool to gain credibility for the program, resulting in stronger connections with possible collaborators/partners.

The evaluation process is central to improving programs by indicating what activities should

be added, eliminated or modified. Evaluation can also be used as an advocacy tool. Sharing information gleaned from an evaluation with all stakeholders can be a springboard for feedback and strategic action, convincing audiences of the value and need for a project, and ultimately generating greater support for a program. The following suggestions are designed to help you present evaluation data effectively:

- Use simple, direct language and highlight key findings.
- Explain what was learned and how and why the program works.
- Be relevant to the audience.
- Tell the program “story” in a compelling way to inspire action.
- Use a variety of presentation techniques including visual displays, written reports, and oral presentations.

Targeting Evaluation Results to Specific Audiences

The most useful evaluations target their results to specific audiences: program staff, foundation boards or trustees, managers, stakeholders, grantees who direct similar or complementary programs, researchers or scientists, and policymakers and other practitioners. Audiences for evaluations have proliferated in the nonprofit world where there is great interest in disseminating lessons learned. All of these audiences have different information needs.

- Grantees’ staff and stakeholders want to know what works in day-to-day operations. Is the program serving the needs of target communities and groups? How can it be improved and focused? What information will convince potential investors to sustain or expand their support? Some grantees may want to use the evaluation process to improve their own management practices.
- Program constituents, i.e., students, parents, and community leaders want information about how well a program served them. Parents want to know if students learned or progressed in school. Students want assurance that their participation in a program will result in higher grades or test scores.
- Foundation management and staff can use evaluation information to design programs, determine investment strategies, and assess the foundation’s overall

effectiveness.

- Foundation boards want to know what the bottom line is: What has changed? What has been accomplished?
- Policymakers, government officials, and community leaders rely on foundation programs to produce innovative solutions to problems. They are interested in learning about new strategies that effectively address community concerns.
- Researchers and scientists look to programs “on the street” to tackle the theoretical questions. Evaluations provide data about solving real-world problems, thereby contributing to the advancement of fundamental science.
- Other stakeholders may be interested in learning how programs serving similar groups work with their clients effectively. What challenges have innovative programs met successfully? How can those procedures be replicated or adapted in new settings?

Using Results to Promote Organizational Learning

Evaluation promotes organizational learning by providing feedback. An evaluation report should:

- Highlight program achievements
- Demonstrate ways a good program can improve educational outcomes
- Encourage other programs in a community to engage in evaluation activities
- Help community stakeholders develop a strategic action plan that targets areas needing improvement and builds upon program success.



