An Evaluation Framework for Community Health Programs

June 2000

An Evaluation Framework for Community Health Programs

Produced by:

The Center for the Advancement of Community Based Public Health

Quinton E. Baker

Denise A. Davis

Ron Gallerani

Victoria Sánchez

Claire Viadro

Based on:

Framework for Program Evaluation in Public Health
(Centers for Disease Control and Prevention)

Community Tool Box: A Framework for Program Evaluation
(University of Kansas)

Preface

At the Center for the Advancement of Community Based Public Health (CA-CBPH) we believe that involving communities as full partners and stakeholders in program development, implementation, and evaluation is vital for building community capacity. Our belief, and commitment to building community capacity, is grounded in principles which hold that: (1) improvements in health are best achieved through the full participation of **communities** in identifying health issues and creating programs to address them; and (2) in **partnerships** among entities such as community-based organizations, health agencies, and universities, each partner has contributions of equal value for achieving goals (e.g., creating healthy communities).

An Evaluation Framework for Community Health Programs can be used by community-based organizations and community health workers to enhance their under-standing of evaluation and to build their capacity to more actively and aggressively participate in evaluation efforts of their programs.

We believe however, that the framework is only one component in building community capacity and that additional efforts are needed to encourage and support more inclusive evaluation processes. Additional efforts can build upon the *Framework's* emphases on the development of measures of program "success" that are relevant to all partners; on considering what is important for communities and other stakeholders to know about their programs; and on producing data that are useful in community settings.

We hope that you will include the *Framework* in your program "toolbox" and that you find application for it in your work. We welcome comments about the *Framework*.

Introduction to the Framework

This document presents a framework that emphasizes program evaluation as a practical and ongoing process that involves program staff, community members, as well as evaluation experts. The overall goal of the framework is to help guide and inform the evaluation process.

The document is **not** a comprehensive manual on how to conduct program evaluation. There are already many excellent resources that meet the technical aspects of program evaluation. Instead, the framework promotes a common understanding of program evaluation. It provides a conceptual roadmap that can be adapted to a variety of settings and within many different groups and communities. The framework may also help individuals or groups with little formal training or experience in program evaluation enter into the evaluation decision-making processes, including consultation with professional evaluators.

The framework is intended to help those involved in program evaluation address the following six questions:

In work we do—especially at the community based organization (CBO) level— evaluation is invaluable, especially if thinking about community based entities or communities being at the center of public health and the work that happens in the public health arena. The ability to have evaluation skills is really key.

> Funding agency
> representative

Guiding Questions for Program Evaluation

1. Who is the evaluation for?
2. What program are we evaluating?
3. What methods will we use in conducting our evaluation?
4. How will we gather and analyze information that is credible and in what forms?
5. How will we justify our conclusions?
6. How can we be assured that what we learn will be used?

The framework also addresses the **quality** of evaluation by asking the question: Will the evaluation be a good evaluation? In evaluation terms, *standards* convey quality and include four key areas to consider:

Four Standards of Program Evaluation

1. Utility (Is the evaluation useful?)
2. Feasibility (Is the evaluation viable and practical?)
3. Propriety (Is the evaluation ethical?)
4. Accuracy (Is the evaluation correct?)

Program evaluators view these four standards as the initial yardstick by which to judge the quality of program evaluation efforts.

How the document is organized

The framework is presented in seven main sections. We present an overview of program evaluation in the first section. We then discuss six steps in program evaluation (illustrated on page 3) in the following sections. These steps address the guiding questions listed on page 1. Although in the real world the steps may not always strictly follow this sequence, as presented each step provides the foundation for the next. A glossary and program evaluation resources are included at the end of the document.

The discussion of each step includes:

◆ **Definition of the step**
◆ **Why the step is important**
◆ **What's involved in completing the step**
◆ **Applying standards**
◆ **Questions (application to your program)**

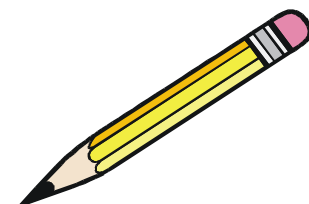**Symbols
to Watch For...**

City of Hope
case example

Key ideas to note
and remember

Questions (application
to your program)

Steps in Program Evaluation

1 Engage Stakeholders

2 Describe the Program

3 Focus the Evaluation Design

4 Gather and Analyze Evidence

5 Justify Conclusions

6 Ensure Use and Share Lessons Learned

STEPS

STANDARDS

Utility

Propriety

Accuracy

Feasibility
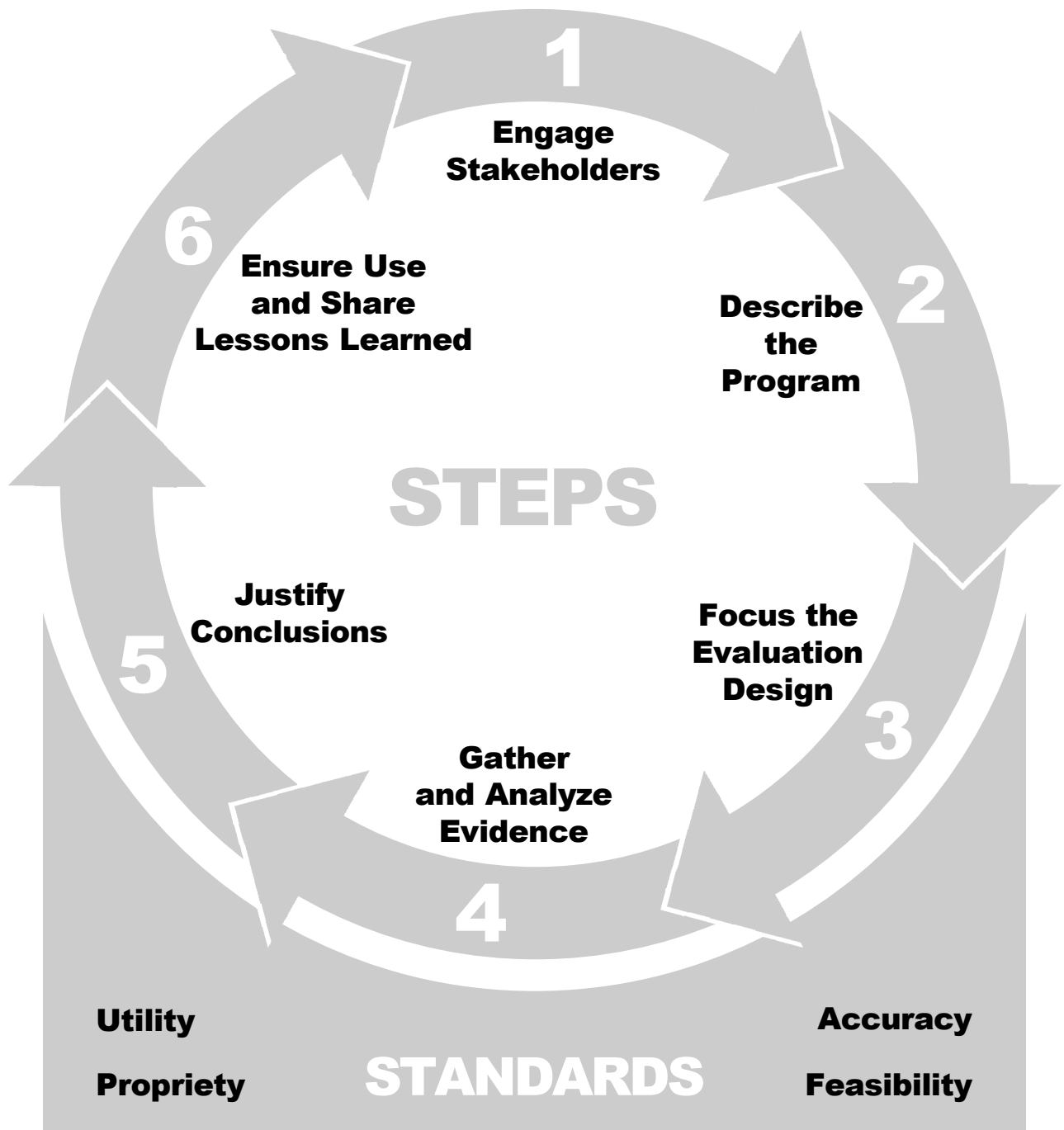
Throughout each step, a case example illustrates the application of the framework in a community program. As just one example, the case is not a rigid or universal illustration of the program evaluation framework. Its main purpose is to illustrate the framework's concepts with a specific health problem, in a particular setting, and with a particular group of people. We invite you to think about how to apply the concepts to your own programs in considering the questions at the end of each section.


City of Hope
Case Example

Overview: The City of Hope received funding from the Centers for Disease Control and Prevention to design and implement a community-based intervention to reduce alcohol injury and death in young men. This funding evolved as an outgrowth of community concern. The problem was documented through a community assessment conducted by community residents, community-based agency representatives, university faculty, and local advocacy groups. This original group believed that achieving this goal would require a long-term commitment of many people and organizations and formalized their commitment by establishing the Partnership to Reduce Alcohol-related Injury and Death (PRAID) coalition.

Problem: Alcohol-related injuries and death in the City of Hope are five times the state rate and two times the rate of cities of similar size and demographics in the region. The assessment revealed a large number of alcohol-related car injuries in two primary locations within a three mile radius of the city. An internal record review by the local police department showed that over a three-year period, 70% of these injuries involved males under the age of 21. A related statistic was the increased number of youth treated for alcohol-related motor vehicle injuries on weekends in the hospital's emergency department. In sum, the
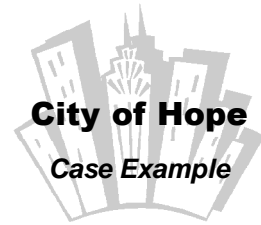
assessment found that a disproportionate number of males between the ages of 15 and 21 were injured or killed during weekend periods and that a high percentage of these injuries and deaths were associated with alcohol-related motor vehicle crashes (MVCs).

Action Plan: The PRAID coalition chose a multi-pronged strategy to reduce alcohol related injury and death experienced by male youth in their community. They proposed to:

1. increase patrols on routes and highways, set up alcohol check points on weekends—specifically on problem routes, and advertise the strict enforcement of legal and monetary penalties associated with drinking and driving (driving while under the influence).
2. promote alcohol education programs in all city schools.
3. require that all students enroll in drivers education with a mandatory three hours on the risks associated with alcohol and driving.
4. reduce the hours for selling liquor in the county.

Evaluation Plan: The PRAID coalition used the six steps presented in this framework to think about the evaluation. Some of the members had some evaluation experience, others had none. They knew however that the evaluation component was important to document their efforts as well as to try to determine if their intervention made a difference.

Overview of Program Evaluation

How many times have you heard yourself say "I'll never do that again!" or, "I'll try this and see how it works out," or, "I really liked that restaurant more than the other"? These examples show how we use evaluation in our everyday lives to make decisions, to test a new idea, or to make comparisons.

When we reflect on our work with communities, we may think of evaluation as a complex, formal, and difficult process. We might equate the evaluation process with the need to collect program information that documents our work or to compare a group that participates in a program with a group that doesn't. Many times the value of evaluation is not immediately apparent to people who work to deliver the actual services (e.g., nurses in a local health department) or to the people who benefit from programs. Sometimes we lose sight of evaluation's benefits when funding agencies emphasize documenting the numbers of people served or when results are not shared with communities because of time pressures to get things done.

However, program evaluation, in fact, does have benefits for people in communities who carry out programs to improve community health. As we developed this document, we spoke to individuals around the country who work daily with community programs. On the following pages, they share their experiences and thoughts on the benefits and applications of program evaluation.

Benefits of Program Evaluation

Reflect on progress; see where we're going and where we're coming from, and improve programs

Evaluation has helped us to focus and study and see where we're at and where we need to go so that we can find our own guns to put forth a better package . . . .

Community nurse

We forget sometimes to look at "Is it really working the way it's already set up?" because we haven't done an evaluation of the program . . . . Do we know it's set up in a way that actually is having change?

Policy analyst

Influence policy makers and funders

The community needs to speak to funders. We need to impact funders . . . . The benefit to evaluation is being able to funnel upward and to trumpet upward, not only be recipients of something.

Community health worker

Build community capacity and engage communities

. . . There's a capacity-building that goes throughout the process, and not that you're going to make researchers out of everybody, but [you're] elevating everybody's understanding and ability to really become engaged.

CBO director

Share what works and what doesn't work with other communities

We want to answer: Does the success of our program have replication in other places, and if so what are the universals there that then can be used as information for program development, planning, evaluation, at other sites?

CBO director

Ensure funding and sustainability

Evaluation is really about people. But it comes through in numbers. And then those numbers, when presented to the person that gives out the funds, it has credibility.

Community nurse

I've worked with a lot of small . . . community based groups that as a result of the evaluation they were able to highlight those things in their annual reports or in their funding requests to different funders and so forth, and as a result of that, be able to get funds. And actually, one of them has been able to double their funding base and also the amount that they get as a result of some very systematic program evaluations we've done together.

Program evaluation consultant

Strengthen accountability

You have so many boards where consumers are members, they have fiduciary responsibilities . . . . You do have people who might not have a lot of training in program design, implementation, evaluation, but ultimately have and are quite aware of having a responsibility for the impact and outcomes of those programs. And so, again, I think that it's imperative that those people have access to evaluate and understand things that they are legally responsible for.

Community health worker

In addition to benefits identified by people who work daily in communities, program planners and evaluators believe program evaluation to be important when there is a need to:

- ◆ collect evidence on a program's effectiveness or impact
- ◆ be accountable to funders, volunteers, staff, and boards
- ◆ identify ways to improve a program:
    - ● assess the needs of individuals, groups, communities
    - ● improve the usefulness of program materials
    - ● determine what works, what doesn't and why
- ◆ clarify program plans
- ◆ improve communication among those involved in the program

In sum, program evaluation is beneficial and important. It is a means to gain feedback and ensure that everyone is working toward successful program implementation. Sometimes informal evaluation is enough. However, when the stakes are high—when a good deal of time or money is involved, when a funder requires it, or when many people may be affected—then it may make sense for your organization to use evaluation procedures that are more formal, visible, and justifiable.

What do we mean by program evaluation?

Thus far we've identified and discussed several aspects of program evaluation, without presenting definitions of the terms. As we define it, a **program** is a series of *activities* supported by a group of *resources* intended to achieve specific *outcomes* among particular individuals, groups, and communities. The term **program** also refers to the effort that is being evaluated. It may apply to *any* action that seeks to improve outcomes for whole communities, for more specific sectors (e.g., schools, work places), or for sub-groups (e.g., youth, people experiencing violence or HIV/AIDS). Examples of different types of programs include:

- ◆ Direct service interventions (e.g., a program that offers free breakfast to improve nutrition for grade school children)
- ◆ Community mobilization efforts (e.g., an effort to organize a boycott of California grapes to improve the economic well-being of farm workers)
- ◆ Research initiatives (e.g., an effort to find out whether disparities in health outcomes based on race can be reduced)
- ◆ Advocacy work (e.g., a campaign to influence the state legislature to pass legislation regarding tobacco control)
- ◆ Training programs (e.g., a job training program to reduce unemployment in urban neighborhoods)

Program planning and implementation are directly related to program evaluation. As you review the framework, it will be helpful to remember that each is part of a larger, integrated program cycle. Thus, while we focus on the program evaluation process here, evaluation is not really separate from the program planning and implementation processes.

Program evaluation is the systematic collection, analysis, and reporting of information about a program to assist in decision-making. Many of us already assess our efforts without necessarily calling it evaluation. We assess the value and impact of our work all the time when we ask questions, consult partners, make assessments based on feedback, and then use those judgments to improve our work.

When we conduct program evaluation, we may answer one or more of the following questions:

- ◆ What have we done?
- ◆ How well have we done it?
- ◆ Whom have we done it to?
- ◆ How much have we done?
- ◆ How effective has our program been?
- ◆ What could we do better or differently?

These questions fit one of three common types of program evaluation: process evaluation, intermediate or short-term (impact evaluation), and longer-term (outcome) evaluation. In general:

Process evaluation answers questions about **how** the program is implemented and how the program outcomes are achieved. It focuses on questions such as: *Is the program being implemented as planned? How is the program achieving its objectives? What activities were conducted? What materials or services did participants receive? What did people experience? How is our coalition working? Do we have the "right" stakeholders?* In addition, process evaluation tracks the strengths and weaknesses of the program and seeks to identify what parts of the program are working and which are not.

Intermediate or short-term evaluation (impact evaluation) answers questions about the **short-term effects** or benefits of a program—as opposed to long-term outcomes such as injury or death. It focuses on questions such as: *What effects did the program have? Can the effects be attributed to the program? Did program participants' knowledge, attitudes, beliefs, or behaviors change as a result of the program? Did the training program achieve its objectives? What happened as a result of the coalition's efforts?* In short, intermediate evaluation addresses the factors that are believed to precede, and that are linked to, longer-term outcomes. In Step 2—Describe the Program—we describe a process (i.e., logic model) that can be used to illustrate the relationship between program actions and outcomes.

Long-term evaluations (outcome) often focus on health status, injury (morbidity), death (mortality), or systems changes. In many health programs the long-term goals are so distant that evaluating them is beyond the range of the specific program evaluation. Outcome evaluation questions are generally related to the overall program goal: *What change in injury or death occurred because of the program? What is the current prevalence (how many cases of "x"*

exist?) or what is the current incidence of "x" (how many new cases of "x" occurred this year?)?

Different professions and academic disciplines use a variety of names to describe the different types of evaluation. Some evaluators consider *formative* evaluation—used to provide ongoing feedback for program improvement—as a sub-set of process evaluation. We agree with them for the purposes of this document. In some evaluation circles, summative evaluation is used to describe both short- and long-term evaluation processes. Some consider program monitoring as another type of evaluation, although in general, its use is limited to measuring program progress.

The importance of evaluation lies in its purpose and not in specific terminology. We encourage you to first think about what you want to find out about your program, how the information will be used, by whom, and for what purposes. Answers to these initial questions will serve to guide your evaluation process before you "name" the type of evaluation you are using.

Additional program evaluation terms used in this document are included in the Glossary. We encourage you to add your own notes or examples to the glossary to build your understanding.

STEP 1

STANDARDS
Utility   Accuracy
Propriety   Feasibility

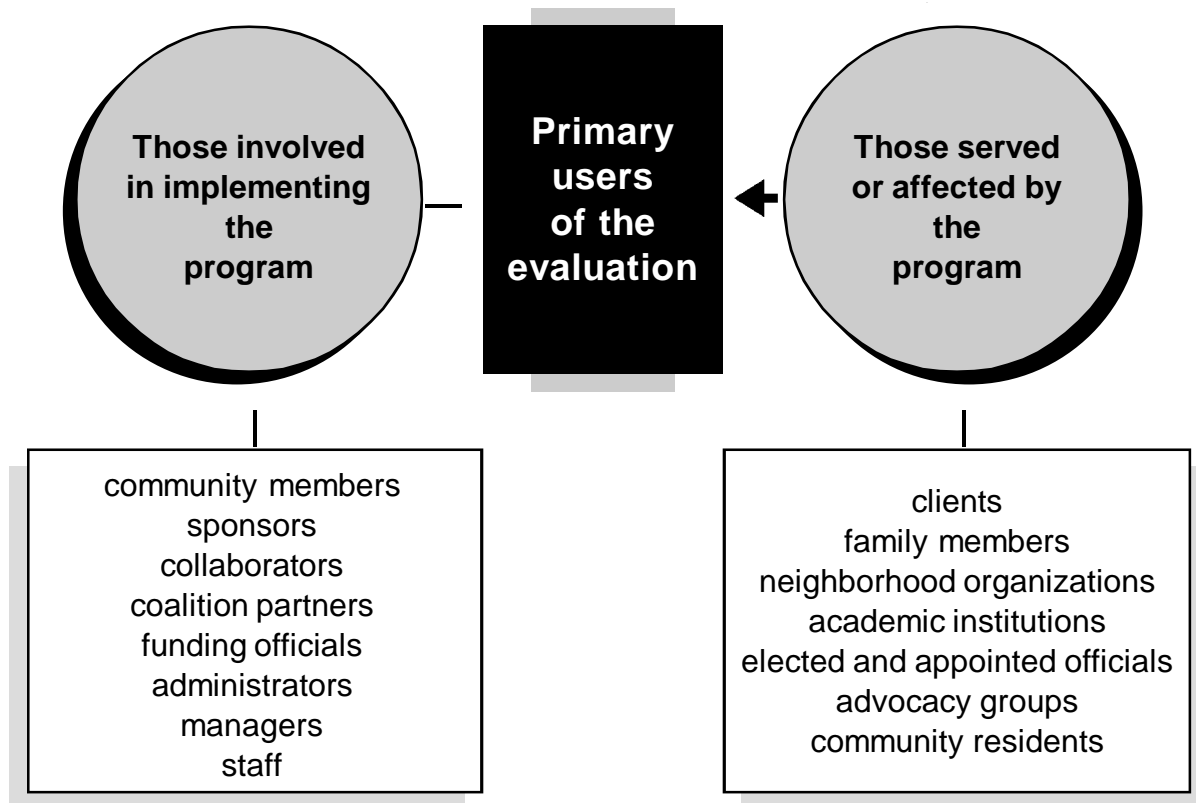# Engage Stakeholders

Definition

Stakeholders are:

- ◆ people who care about what will be learned from the evaluation and about what will be done with the knowledge gained

Community people, health agency people, and representatives from community-based organizations have described stakeholders as:

- ◆ people with interest in the program
- ◆ all the necessary partners
- ◆ collaborators
- ◆ those affected by the program
- ◆ shareholders
- ◆ all who have the same shared interest

Another way of thinking about evaluation stakeholders is to think about the type of stakeholder. Individuals and groups from many different arenas can be categorized in three main groups of stakeholders, depending on what their role is in the planning and evaluation process. The following diagram illustrates these three groups.

Types of Stakeholders

Those involved in implementing the program	Primary users of the evaluation	Those served or affected by the program

community members sponsors collaborators coalition partners funding officials administrators managers staff	clients family members neighborhood organizations academic institutions elected and appointed officials advocacy groups community residents

The primary users of the evaluation are made up of people from the two other groups. The primary users are the specific individuals or groups who are in a position to decide about and/or do something with the results. A successful evaluation will designate its primary intended users—such as community based organizations, groups of citizens, program staff and funders—early in the evaluation's development, and will maintain frequent interaction with users to be sure that the evaluation specifically addresses their values and needs.

Why is it important to include different stakeholders?

Evaluation cannot be done in isolation. Almost everything done in community health and development work involves partnerships—alliances among different organizations, board members, those affected by the problem, and others. Therefore, any serious effort to

evaluate a program must consider the different values held by the partners or stakeholders.

Stakeholders must be part of the evaluation so that their unique perspectives are understood. When stakeholders are not appropriately involved, evaluation findings may be ignored, criticized, or resisted.

What's involved in identifying stakeholders?

When identifying stakeholders we ask:
◆ Who cares about the program?
◆ What do they care about?
◆ Which individuals support the program?
◆ Which individuals are openly skeptical of, or antagonistic toward, the program?

Opening an evaluation to opposing perspectives and enlisting the help of potential program opponents can strengthen the evaluation's credibility. Likewise, individuals or groups who could be adversely or inadvertently affected by changes arising from the evaluation have a right to be involved. This means including those who would be affected if program services were expanded, altered, limited, or ended because of the evaluation.

The amount and type of stakeholder involvement will be different for each program evaluation. In many instances stakeholders will be directly involved in designing and conducting the evaluation. They can be kept informed about progress of the evaluation through periodic meetings, reports, and other means of communication.

Some community initiatives form an evaluation team—made up of various stakeholders—as part of their overall planning and evaluation process. While not all members have program evaluation experience or training, at least some members do. This type of arrangement can work well when there is a commitment to build the capacity of community-based organizations, local advocacy groups, or interested citizens to learn about and participate in program evaluation.

Early on, an informal project planning committee came together to plan the program and the evaluation. Because they knew how important it was to include a variety of stakeholders, they identified and recruited community stakeholders (parents, students, and local business owners) in the initial phases of the project through public forums, by word of mouth, and as they conducted their community assessment. As these stakeholders—who became members of the PRAID coalition—met and shared their ideas, the following interests emerged:

Identifying Stakeholders

Early Stakeholders	What They Cared About Most
Mothers Against Drunk Driving (MADD)	Reducing death and injury among young people
Students Against Drunk Driving (SADD)	Preventing alcohol use in teens, especially related to motor vehicle crashes (MVCs)
Health department	Protecting residents from alcohol-related MVCs
Hospital	Reducing the resources spent on MVCs
High schools	Changing teens' positive image of drinking
Police department	Reducing everyone's risk on the highways
City officials	Changing the town's image
School of Public Health	Improving the health of the community

After the project staff were on board, staff spent many hours with stakeholders—in meetings or one-on-one talks—to understand what the stakeholders cared about and to find ways to share concerns with other stakeholders. They relied on formal and informal communication to make sure that as many stakeholder voices were heard as possible.

As the project matured, new stakeholders became involved, including the restaurant association president and a representative from the regional beverage (alcohol) distributor. The latter were primarily interested in balancing the negative image of drinking and driving among teens with responsible drinking in legal drinkers (e.g., adults). Their views were important as the PRAID coalition refined its program plan and focused the evaluation.

Applying standards

Including stakeholders in evaluation planning and implementation is one way to ensure a quality evaluation to meet evaluation standards. The two standards that apply most directly to Step 1—Engage Stakeholders—are **utility** and **propriety**. As you carry out this step, the questions presented in the following table will help you to clarify and achieve these standards.

Standards for Step 1: Engage Stakeholders

Questions	Standards
◆ Have you included individuals and organizations that will be affected by the evaluation in your evaluation planning group? ◆ Have you considered adding new stakeholders as your program evaluation is implemented? ◆ Are participants in the evaluation planning group trustworthy and competent?	**Utility:** Ensures that the evaluation is useful and answers questions that are directly relevant to users.
◆ Are individuals clear about what is to be done, how, by whom, and when? ◆ Is there a written understanding? ◆ Have steps been taken to assure that all stakeholders and the population served will be respected and their values honored? ◆ Have conflicts of interest been discussed to ensure that the results or findings will not be compromised?	**Propriety:** Ensures that the evaluation is an ethical one, conducted with regard for the rights and interests of those involved.

In summary, Step 1—Engage Stakeholders— represents a process through which many voices are heard. As the first step, it makes the benefits of the evaluation clear to all stakeholders. Completing this step helps ensure that the focus of the evaluation—and ultimately the results of the evaluation—supports the needs of the stakeholders.

Think about a program evaluation in which you are a stakeholder or a program for which you would be responsible for identifying and bringing people together. The following questions may help you think about how to approach Step1—Engage Stakeholders.
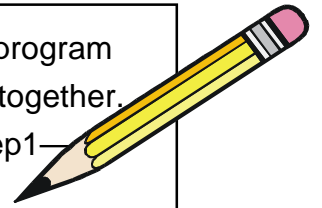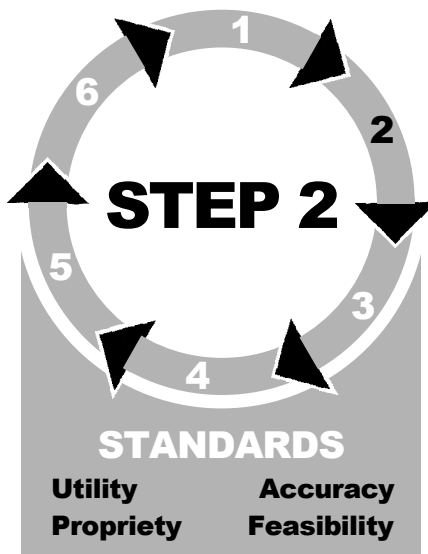
Who are the people who fall into the three types of stakeholder categories, i.e., those involved in implementing the program, those served or affected by the program, and the primary users of the evaluation?

How could you go about finding out what each stakeholder cares about?

What communication strategies could you use to ensure that different interests are represented?

What challenges or barriers might you face in identifying and recruiting stakeholders?

How could you deal with these challenges or barriers?

Describe the Program

STANDARDS

Utility   **Accuracy**
Propriety   **Feasibility**

Definition

A **program description**:

- summarizes the program being evaluated

- explains what the program is trying to accomplish and how it tries to bring about those changes

- illustrates the program's core components

- establishes a program's ability to make changes

- specifies its stage of development

- describes how the program fits into the larger organizational and community environment

Why is it important to describe the program?

How a program is described sets the frame of reference for all future decisions about its evaluation. For example, one program may be described as "attempting to strengthen enforcement of existing laws that discourage underage drinking." Another program may be described as "a program to reduce drunk driving by teens." The first program specifies enforcement, where the second program is more broadly defined as reducing drunk driving. The

different emphases of these two programs will shape the direction of evaluation efforts.

In addition, the program description allows members of the evaluation group to compare the program to similar efforts, and makes it easier to figure out what parts of the program brought about what effects.

Different stakeholders may have different ideas about what the program is supposed to achieve and why. For example, in a program to reduce teen pregnancy some stakeholders may believe this means increasing access to contraceptives, while others may believe it means focusing on abstinence. Evaluations done without agreement on the program definition aren't likely to be very useful. In many cases, the process of working with stakeholders to develop a clear and logical program description will bring benefits long before data are available to measure program effectiveness. (This is a good example of why it's important to identify stakeholders [Step 1] before proceeding with the next steps.)

As we will see in the next step, PRAID coalition members decided to focus on reducing injuries and death in young men—a broad goal. Coalition members had different ideas, however, about how to reduce the injury and death rates. Police department representatives were sure that increased penalties for drinking and driving were the key to reducing injury and death; school administrators were convinced that education was key. Through a participatory planning process, coalition members and other stakeholders developed a long-term plan that incorporated different ideas and priorities within the broad goal of reducing alcohol-related injuries and deaths among young men. Their program description spelled out their plan to achieve this goal.

Program Description

What's involved in describing, or what are the elements of, a program?

There are seven elements of a program description. A well defined program description lays the foundation for focusing the evaluation (Step 3).

1. **Statement of need**
2. **Expectations**
3. **Activities**
4. **Resources**
5. **Logic model**
6. **Context**
7. **Stage of development**

The first five elements are related in a very concrete way to the development of the program. The last two elements—the context and stage of program development—are important, yet are larger than the program itself and are outside the technical description of the program. The first four elements are shown in the following illustration.

Four elements in describing a program

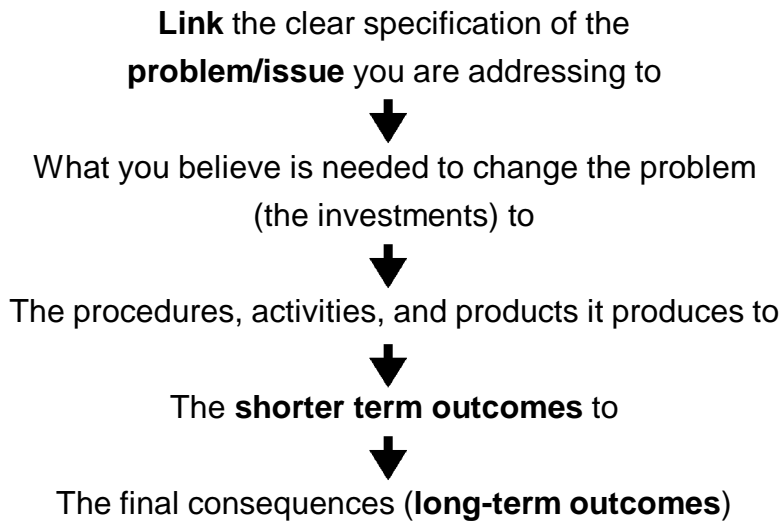| Statement of Need	Expectations	Activities	Resources
What we need to know to describe the problem/issue	What our results will be	What we need to do to change the problem	What resources we need
Include: Who is affected? How big is the problem? Is it changing? How is it changing?	**Define:** What are our expectations? What are immediate, intermediate, and long-term consequences? What are the objectives, goals, mission and vision of our program?	**Identify:** What are the specific strategies and actions we need to take?	**Determine:** ◆ time ◆ talent ◆ equipment ◆ information ◆ money ◆ other assets

Case Example: Elements of the City of Hope's program to reduce injury and death in male youth

Statement of Need (the problem): Three-year increase in the number of alcohol-related motor vehicle crashes (MVCs) among males under 21

Goal: Reduce alcohol-related injury and death

Objective: By 2005 reduce the incidence of alcohol-related MVCs in males under 21 by 50%.

Expectations	Activities	Resources
Sub-Objective A:		
Increase legal sanctions and enforcement for driving under the influence (DUI)	◆ Increase highway patrols/enforcement in the county (e.g., alcohol checkpoints) ◆ Increase DUI sanctions for males under 21	◆ Supplemental funding for checkpoints ◆ Training for law enforcement officials
Sub-Objective B:		
Increase knowledge among underage youth of the risks associated with drinking and driving	◆ Increase the participation of young men in alcohol education programs and driver's education	◆ Alcohol education curricula ◆ Funding for alcohol and drivers' education programs
Sub-Objective C:		
Eliminate the availability of alcohol to males under 21	◆ Reduce the hours for selling liquor in the county ◆ Increase legal and financial penalties for alcohol vendors selling to males under 21	◆ Staff to perform "sting" operations to monitor vendors ◆ Resources to disseminate information about the program to alcohol outlets
Sub-Objective D:		
Increase collaborative problem-solving among community, agency, and university stakeholders	◆ Build and sustain coalition of interested parties	◆ Community groups, professional organizations and municipal agencies ◆ Program staff

A **logic model** is a map that graphically illustrates how your program activities will lead to the expected short-term— and long-term—outcomes. A logic model will make it clear whether your program makes sense and whether the expected outcomes are realistic given the planned activities. In short, the model shows what your program is expected to achieve and how it is expected to work, based on an expected chain of events that:

Needs, expectations, and activities are reflected in the development of the Logic Model.

Link the clear specification of the
problem/issue you are addressing to

⬇

What you believe is needed to change the problem
(the investments) to

⬇

The procedures, activities, and products it produces to

⬇

The **shorter term outcomes** to

⬇

The final consequences (**long-term outcomes**)

(Adapted from University of Wisconsin *Evaluating Collaboratives: Reaching the Potential*, July 1999, G3658-8)

The logic model synthesizes the main program elements into a picture of how the program is supposed to work. Often displayed in a flow chart, map, or table, the logic model portrays the sequence of steps leading to program results. A simple logic model might look like this:

Problem or issue	Investments	Procedures, activities, and products	Short-term outcomes (related to sub-objectives)	Longer-term outcomes (related to program goal)
Pregnant teens are not using prenatal services	*If:* There is an investment of time and money	to develop a resource directory	*then:* teens will be informed about available services	*then:* teens will be able to gain access to the services to meet their needs

Logic models have several benefits. They:

- allow stakeholders to improve and focus program direction
- reveal assumptions about conditions for program effectiveness
- provide a frame of reference for one or more evaluations of the program
- can be a basis for estimating the program's effect on endpoints that are not directly measured

A simple model for the program is illustrated in the following:

Logic Model

Problem or issue	Investments	Procedures, activities, and products	Short-term outcomes (related to sub-objectives)	Longer-term outcomes (related to program goal)
Alcohol-related injury and death in males under 21	*If:* ◆ law enforcement resources are committed ◆ educational resources and skills are committed ◆ people come together	◆ to strengthen enforcement of DUI laws ◆ to provide alcohol and driver's education ◆ to foster change in alcohol sales	*then* *we'll see:* ◆ increased citations ◆ increased knowledge about risks ◆ decreased sales ◆ changes in drinking and driving behaviors	*then:* there will be a reduction in injuries and deaths due to alcohol

In "story form" the logic model describes a program that says: "When law enforcement resources are committed, where educational resources and skills are committed, and when people come together, there will be stronger law enforcement, increased provision

of education services, and efforts to change alcohol sales policies. When these activities are in place, we will see an increase in penalties for driving under the influence, increased evidence of knowledge about alcohol-related risks, and a decrease in alcohol sales to underage youth. These changes will ultimately lead to a reduction in injuries and deaths related to alcohol in the under-21 male population in our area."

A description of a program's **context** considers the important features of the environment in which the program operates. This includes understanding the context in terms of:

- history
- politics
- geography
- social and economic conditions
- what other organizations have done

A realistic and responsive evaluation will be sensitive to a broad range of potential influences on the program. An understanding of the context lets users interpret findings accurately and assess the findings' generalizability. For example, a program to improve housing in an inner-city neighborhood might be a tremendous success, but would likely not work in a small town on the other side of the country without specific adaptation.

When PRAID coalition members came together to discuss how they could reduce alcohol-related injuries, they first talked about past efforts. The local MADD chapter had successfully advocated for new alcohol policies at the state level but had had little success with the county commission in the past. Two new county commissioners, however, had recently been elected; one of them had lost a nephew in an

Program Context

25

alcohol-related motor vehicle crash—a potential ally for their initiative. Stakeholders also discussed the local DARE program and how it had been part of alcohol prevention programming for younger children for several years. The PRAID coalition knew that the police department was heavily invested in the DARE program, raising concerns that resources to fund the proposed PRAID strategies might be perceived as competing with existing programs such as DARE. Finally, the coalition members thought about how their program's educational components could be made relevant to a wide variety of students—acknowledging cultural differences, and also recognizing that individuals learn in many different ways. As the PRAID stakeholders refined their program and the evaluation, they were sensitive to these issues.

A program's **stage of development** also affects the evaluation process. For example, an evaluation of a new program may differ significantly from an evaluation of a program that has existed for a number of years. One way of viewing the intent of evaluation at different stages is through the different goals of evaluation at three commonly recognized stages:

| Stage of Program	Evaluation Goal
Planning: program activities are untested	To refine plans as much as possible
Implementation: program activities are being field tested and modified	To see what happens in the "real world" and to improve operations
Effects or outcomes: enough time has passed for the program's effects to emerge	To identify and understand the program's results, including those that were unintentional

Applying standards

As in the previous step, you can assure that the evaluation is a quality one by incorporating standards into the program description. The two standards that apply most directly to Step 2—Describe the Program—are **accuracy** and **propriety**. As you carry out this step, the questions presented in the following table can help you to clarify and achieve these standards.

Standards for Step 2: Describe the Program

Questions	Standards
◆ Is the program description complete? ◆ Have you documented the context of the program so that likely influences on the program can be identified?	**Accuracy:** Ensures that the findings can be considered correct.
◆ Is the evaluation complete and fair in assessing all aspects of the program, including its strengths and weaknesses?	**Propriety:** Ensures that the evaluation is an ethical one, conducted with regard for the rights and interests of those involved.

In summary, Step 2—Describe the Program—outlines the foundation to ensure a well thought out program evaluation. It is important to keep in mind that the breadth and depth of a program description will vary for each program evaluation, meaning that many different activities may be part of developing the description. Multiple sources of information may be pulled together to construct a well-rounded description. Discussion with stakeholders can confirm the accuracy of an existing program description. Descriptions of what's going on may be checked against direct observation of activities in the field. Addressing contextual factors (such as staff turnover, inadequate resources, political pressures, or strong community participation) that may affect the program can broaden an otherwise narrow program description.

Think about a program evaluation in which you are involved. The following questions may help you think about how to approach Step 2—Describe the Program.
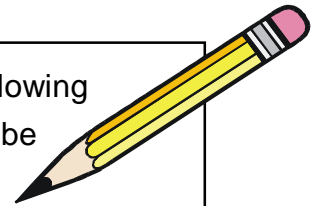
How are your program's goals, objectives, and strategies defined?

How are your program's activities, processes, and products linked to the program's outcomes?

What resources might be available to implement the program?

What else might be happening in your community that could have an impact on your program? What other programs have been tried, and by whom?

Is your program new or has it existed for a year or more? (As you read the next step, think about how this could influence how you focus your evaluation.)
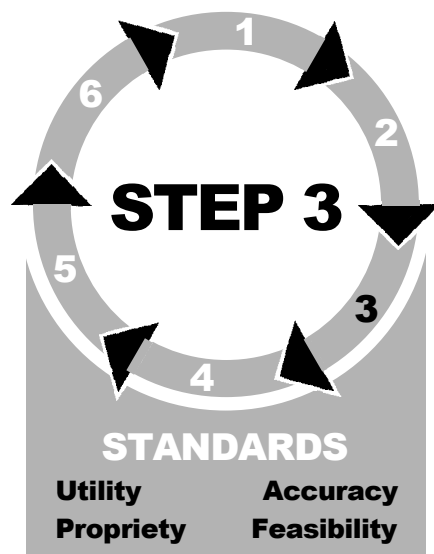
Focus the Evaluation Design

Definition

> **Focusing the evaluation design** means:
>
> ◆ carrying out advance planning about where the evaluation is headed and what steps it will take to get there
>
> ◆ developing a well-focused plan or strategy to improve the usefulness of the evaluation to intended audiences

Why is it important to focus the evaluation design?

This question involves determining the users and the uses of the evaluation. It gets at the direct purpose of the evaluation. Some common purposes could be to:

- ◆ test program effectiveness
- ◆ make a case to change program practices
- ◆ assess the effects of a program on a specific population
- ◆ justify continued financial or political support

Additional considerations include making sure that the chosen evaluation approach answers stakeholder questions and that the process provides continuous feedback at all levels of program operation.

It's a stipulation within the grant to have an evaluation plan in place from the start. [It's] necessary that periodic reports are submitted to the funder on where the evaluation stands in the process.

Community Member

What's involved in focusing an evaluation design?

Once you have determined who or what the evaluation is for, there are several steps to focus the evaluation design. They include:

- ◆ determining the information needs of your various stakeholders
- ◆ assessing the best techniques to describe and measure your program activities
- ◆ assessing what qualitative and quantitative data may be available to you
- ◆ determining the design method that best answers the key questions set by stakeholders
- ◆ preparing a written agreement that summarizes the evaluation procedures and specifies the roles and responsibilities of all involved

There are seven basic issues to consider when focusing an evaluation:

1. Users	3. Purpose	5. Design	6. Methods
2. Uses	4. Questions		7. Agreements
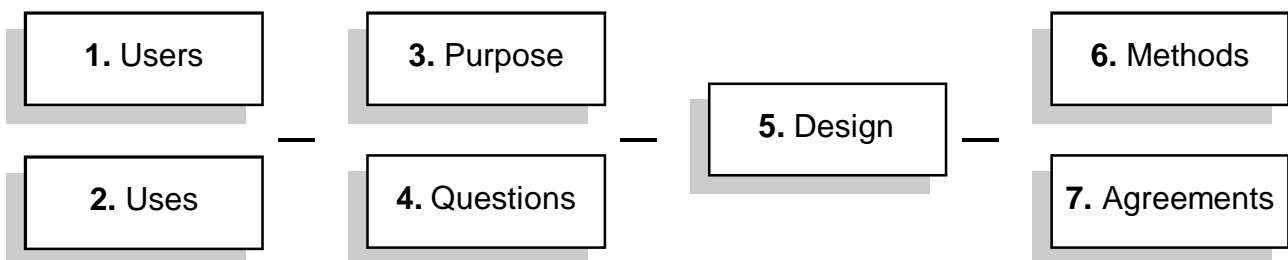
Users are the specific persons who will use evaluation findings. Because they directly experience the consequences of the trade-offs that are part of any evaluation, they have a right to participate in choosing the focus for the evaluation. When users are encouraged to clarify intended uses and identify priority questions and preferred methods, the evaluation is more likely to focus on things that will inform and influence future actions.

Uses describe what will be done with what is learned from the evaluation. Information collected may have varying uses, which should be described in detail when designing the evaluation. This directly feeds into defining the purpose of the evaluation and the questions to be asked.

Illustrated here is a list of evaluation users and their intended uses for the information:

Users	Uses
Police Department	Could use the information from the evaluation to determine how to allocate their internal resources to better serve the public.
PRAID Coalition	Could employ evaluation information to bolster data required for additional funding for program sustainability.
City Officials	Could use evaluation information to formulate new legislation related to each objective.

Purpose refers to the general intent of the evaluation. A clear purpose serves as the basis for the evaluation questions, design, and methods. Community organizations might become involved in evaluation to:
- ◆ gain new knowledge about program activities
- ◆ improve or fine-tune existing program operations (e.g., program processes or strategies)
- ◆ determine the effects of a program by providing evidence concerning the program's contributions to a long-term goal
- ◆ affect program participants by acting as a catalyst for self-directed change (e.g., teaching

[Evaluation] is demanded when we work with external organizations. There's a demand to have an evaluation conducted and be able to have, you know, meaningful quantitative data to show results.

Community Member

evaluation skills to staff and stakeholders)

◆ build capacity or the ability to develop trusting working relationships among coalition members, health care providers and local government representatives

The primary purpose of the evaluation as proposed by the PRAID coalition was to determine whether the stated program goals could be achieved by the year 2005. The coalition wished to gain insight as to whether the new approaches they had outlined in the form of specific objectives would have any measureable effect on alcohol-related injury and death in young men. In addition, coalition members were interested in the processes by which the primary goal would be achieved. They viewed the building of relationships among PRAID coalition members, community members, service providers, and local government representatives as key to the project's ability to effectively implement its action plan. Recognizing that the implementation of specific program objectives had the potential to build community resources and skills as well as community capacity, coalition members decided that the evaluation should also focus on measuring and describing changes in these community processes.

Purpose

Questions establish the aspects of the program that will be addressed and encourage stakeholders to reveal the questions they believe the evaluation should answer. Clearly worded questions that address changes in individuals, within organizations, or in whole counties (units of analysis) will help guide method selection and evidence gathering.

The questions that the PRAID coalition members identified were strategically linked to the program objectives to reveal what they believed the evaluation should address.

Questions

Program Objectives	Evaluation Questions	Intervention Strategies	Short-term Outcomes	Long-term Outcomes
Increase legal sanctions for driving under the influence (DUI)	Will legal sanctions and increased high-way patrols reduce the incidence of DUI behavior in males under 21?	◆ Increase patrols and enforcement (e.g., alcohol checkpoints) ◆ Increase DUI sanctions for males under 21	Increased citations	Reduction in alcohol-related motor vehicle crashes
Increase knowledge among males under 21 about the risks associated with drinking and driving	Will increased participation of males under 21 in alcohol and driver's education result in lower alcohol-related morbidity and mortality?	Increased participation of males in alcohol education and driver's education programs	◆ Increased knowledge about risks associated with alcohol ◆ Changes in drinking and driving behaviors	Reduction in injury and death related to alcohol in the under-21 male population
Limit the availability of alcohol to males under 21	Will increased legal penalties and time of day restrictions on the sale of alcohol to males under 21 result in less underage drinking and driving?	◆ Reduce hours of liquor selling establish-ments in the country ◆ Increase legal and financial penalties for alcohol vendors sell-ing to males under 21	Decreased sales to males under 21	Reduction in injury and death related to alcohol in the under-21 male population

33

Program Objectives	Evaluation Questions	Intervention Strategies	Short-term Outcomes	Long-term Outcomes
Increase citizen participation; develop leader-ship, resources, and social and interorganiza-tional skills	Will the process of working collectively toward the achievement of the program goal develop community capacity?	◆ Regular meetings of the PRAID coalition subgroups to assist in program planning, implementa-tion and evaluation	Increased sense of community; improved social integration	Improved community problem-solving

Design refers to how the evaluation's questions, methods, and overall processes are constructed. The evaluation should:

- be organized from the start to achieve specific agreed-upon uses
- have a clear purpose that is focused on the utility of what is learned
- help those who carry out the evaluation to know *who* will do *what* with the findings

The process of developing a clear design can highlight the ways that stakeholders, through their many contributions, can improve the evaluation and facilitate the use of the results.

When choosing evaluation **methods**, evaluation planning groups usually consider whether the methods will allow stakeholders to effectively answer the questions that are being raised. Information on how the evaluation data will be used is key to decisions about methods. Different evaluation questions require different kinds of data that, in turn, require different types of methods. Counting how many times an event occurs calls for *quantitative* methods, whereas *qualitative* methods might be used to develop an under-standing of the social context in which an event occurs.

Different evaluation designs can require a variety of methods, including the following:

- ◆ systematic participant observation
- ◆ structured or semi-structured interviews
- ◆ focus groups
- ◆ descriptive or analytical surveys

The PRAID coalition was interested in reducing alcohol-related injury and death among the community's male youth. In selecting a design to evaluate the interventions' longer-term effects, the coalition determined that there were no nearby communities of like size and demographics suitable to serve as a comparison group. Taking into account resources and the evaluation questions of interest to stakeholders, the coalition decided that their strategy would include assessing trends over time, identifying key issues, and determining the relationships among key events and outcomes. The PRAID coalition also considered testing the effectiveness of the alcohol and driver's education programs by comparing males under 21 in the City of Hope with males under 21 in a like-sized city who did not get the training. Due to financial constraints and political considerations, however, this approach was not considered feasible.

Design and Methods

Agreements summarize the evaluation procedures, clarify everyone's role and responsibilities, and describe how the evaluation procedures will be implemented. Elements of an agreement include: statements concerning the intended users, uses, purpose, questions, design, and methods as well as a summary of the deliverables, timeline and budget. An agreement might be a legal contract, a memorandum of understanding, or a detailed protocol. Creating an agreement establishes a mutual understanding of the activities associated with the evaluation. It also provides a basis for modification if necessary.

Applying standards

   As we have noted in previous steps, you can help ensure the quality of your evaluation by considering relevant evaluation standards throughout the evaluation process. The standards that most directly apply to Step 3—Focus the Evaluation Design—are **utility**, **feasibility**, **propriety**, and **accuracy**. As you participate in focusing the evaluation design, the questions presented in the following table can help you to clarify and achieve these standards.

Standards for Step 3: Focus the Evaluation Design

Questions	Standards
◆ How can the evaluation be planned, conducted, and reported so that it, in turn, encourages use by stakeholders?	**Utility:** Ensures that the evaluation is useful and answers questions that are directly relevant to users.
◆ Are the evaluation procedures practical? Will they keep disruption of daily activities to a minimum? ◆ Have you considered the political interests and needs of various groups in planning the evaluation? ◆ Have you assessed the costs of technical resources and time?	**Feasibility:** Ensures that the evaluation makes sense, takes into consideration the interests of various groups, and can be cost effective.
◆ Have you addressed issues of conflict openly and honestly? ◆ Have you planned and implemented sound, ethical consistant procedures to ensure findings are correct?	**Propriety:** Ensures that the evaluation is an ethical one, conducted with regard for the rights and interests of those involved.
◆ Have you described the purposes and procedures of the evaluation in detail? ◆ Can the purposes and procedures be indentified and assessed?	**Accuracy:** Ensures that the findings are considered correct.

In summary, Step 3—Focus the Evaluation Design— represents a process through which a design is structured to capture the information that all stakeholders within the effort agree are critical. It ensures that the evaluation design meets the needs of all users, that the process answers the questions that have been raised, and that the evaluation takes into account constraints imposed by time and the availability of technical resources.

Think about a program evaluation in which you are involved. The following questions may help you approach how you think about Step 3—Focus the Evaluation Design.
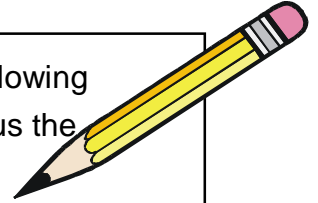
**What kinds of information needs are your users likely to have?**

How could the information produced by the evaluation be used?

What are possible evaluation questions for your program?

What types of evidence could help you show that the program had the intended effect?

How could you go about getting the evidence you need?

Have you thought about who could help you with the technical and design aspects of your evaluation?
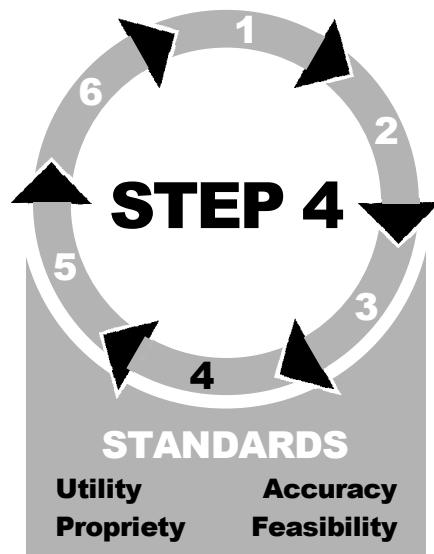
Gather and Analyze the Evidence

STANDARDS

Utility   **Accuracy**
Propriety   **Feasibility**

Definition

> **Gathering and analyzing evidence** means:
>
> ◆ assembling:
>
>   ● the raw material of a good evaluation
>   ● information that gives a well-rounded picture of the program
>   ● those data that have been analyzed and synthesized
>
> ◆ presenting meaningful results

Why is it important to gather credible evidence?

Stakeholders should view the information gathered as believable, trustworthy, and relevant to their questions. Credibility is based on the questions asked at the beginning of the evaluation process and stakeholders' motives for asking them. In other words, standards of credibility depend on the questions asked.

Having credible evidence strengthens evaluation results as well as the recommendations that follow from them. Recognizing that all types of data have limitations, you can strengthen the credibility of an evaluation design by using multiple procedures for gathering, analyzing and interpreting data. Increased up-front participation by stakeholders also enhances credibility because they will be more likely to

accept the evaluation's conclusions and act on its recommendations.

In some situations, you may need to ask evaluation specialists for advice on the most appropriate method(s) to use, given the data restrictions, resource constraints and standards of evidence set by stakeholders. This is especially important in situations where concern for data quality is high or where errors of inference would have serious consequences.

An essential part of a good evaluation is a review of what is known, what has been done before and what has been done elsewhere. Reviewing published literature and unpublished documents can help strengthen the process of designing an evaluation. Program participants and other stakeholders also can be important sources of background information. Familiarity with other evaluations or research on similar interventions (or different interventions designed to address the same problem) can help you develop criteria for judging your own program.

What's involved in gathering and analyzing evidence?

The following features of evidence gathering typically affect how credible an evaluation is seen as being:

1. **Indicators**
2. **Sources of evidence**
3. **Quality**
4. **Quantity**
5. **Logistics**
6. **Analysis and synthesis**
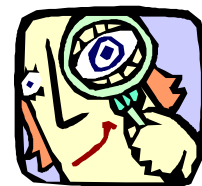
Indicators are used to judge programs. An indicator is a category of change. Indicators translate general concepts about a program and its expected effects into specific, measureable parts. They should reflect the aspects of the program that are most meaningful to monitor. Several indicators are usually needed to track the implementation and effects of a complex program or intervention.

Examples of indicators include:

Category of Change	Examples of Measures
Service delivery improvement (a measure of program capacity)	◆ Client satisfaction level ◆ Number of patients immunized
Changes in individual knowledge, skills, behavior	◆ Participant behavior
Changes in population health status	◆ Injury rates ◆ Death rates
Changes in the environment	◆ Programs, policies and practices

Logic models (see Step 2) can help develop indicators. A logic model can be used as a template to define a full spectrum of indicators along the pathway that leads from program activities to the program's expected effects. Both qualitative and quantitative indicators can be developed.

Sources of evidence in an evaluation may be people, documents, or observations. More than one source may be used to gather evidence for a given indicator. In fact, selecting multiple sources of evidence provides an opportunity to include different perspectives about the program, thereby enhancing the evaluation's credibility. For instance, program documentation reflecting an internal (staff) perspective could be combined with key informant interviews with program users. Mixing perspectives provides a more comprehensive view of the program. In addition, the integration of qualitative and quantitative information can yield evidence that is more complete and useful, meeting the needs and expectations of a wider range of stakeholders.

Selected Sources of Evidence

◆ Written surveys
◆ Personal interviews
◆ Observation
◆ Document analysis
◆ Case studies
◆ Logs, activity forms, registries

◆ Group assessments
◆ Diaries or journals
◆ Geographical mapping
◆ Expert or peer review
◆ Concept mapping

The PRAID coalition identified a variety of possible indicators that related to the program's objectives. Possible sources of evidence included police records, emergency room data, questionnaire data from a community-level survey to assess and monitor the extent of the problem, new community assessment data, key informant interviews with youth attending the local high schools, and so forth. Coalition members were also interested in using standardized measures that would make it possible to measure changes over time and across sites. Finally, to assess program progress coalition members identified process indicators and sources of evidence to document short-term outcomes.

Indicators and Sources of Evidence

Objectives	Indicators	Sources of Evidence
Increase legal sanctions for DUI (driving under the influence) for males under 21	◆ Number of DUI charges for underage youth	◆ Police records of violations ◆ Juvenile court filings
Use driver's education and alcohol education programs to increase knowledge among young men about the risks associated with drinking and driving	◆ Level of participation in driver's education and alcohol education programs	◆ School records for attendance at driver's education and alcohol education programs
Limit the availability of alcohol to males under 21 through increased legal and financial penalties for alcohol sales to this group	◆ Number of on-site inspections of community liquor vendors ◆ Liquor licenses on file	◆ Violations records of liquor vendors ◆ Inventory of liquor-serving establishments

Quality refers to the appropriateness and integrity of information gathered in an evaluation. Well-defined indicators make it easier to collect high quality data that are reliable and informative. Other factors that affect quality include:

- ◆ instrument design
- ◆ data collection procedures
- ◆ training of those involved in data collection
- ◆ source selection
- ◆ coding
- ◆ data management
- ◆ routine error checking

Obtaining quality data involves trade-offs (i.e., breadth vs. depth). Thus, stakeholders must decide at the beginning of the evaluation process what is most important. A practical evaluation should strive for a level of quality that will match stakeholders' standards for credibility.

> In determining what was most important to them, the PRAID coalition decided that uniform instruments (unchanged or standardized) should be used to improve the quality of the evaluation's survey data. Thus, the same data collection instrument was used at each data collection site. In addition, the coalition developed an interview guide for key informant interviews. Other steps taken to assure quality included training interviewers. Training was also provided for individuals involved in data collection, coding, and data management. The training emphasized standardized recordkeeping to ensure the completeness of data. Unexpected changes in one agency's recordkeeping practices required the coalition to reconsider whether and how the agency's records could be used; the coalition also considered substitute sources of evidence that might be available. Coalition members were careful

Quality

to pick sources of data that were fairly complete as well as accessible (e.g., police blotter records). All these actions served to reduce the amount of error within the data and improved the data's overall reliability.

Quantity refers to the amount of evidence gathered in an evaluation. Evaluation stakeholders should estimate in advance the amount of information that will be required, and establish criteria to decide when to stop collecting data. Quantity affects the level of confidence or precision users can have (i.e., how sure are we that what we've learned is true?). It also partly determines whether the evaluation will be able to detect program effects. All evidence collected should have a clear, anticipated use.

PRAID coalition members decided to use both qualitative and quantitative data to develop as wide a picture as possible of the program. First, they decided to use key informant interviews to elicit information from the under-21 population. After a series of interviews had been conducted, the interviewers began reaching a "saturation" point in data collection, where different key informants began to express similar patterns or themes. Second, the coalition collected police, juvenile court, and hospital data for the three-year period before the intervention, and annually throughout the intervention. Third, the coalition's university partners helped other stake-holders determine the quantitative evidence required to assess the program's effectiveness. Specifically, the coalition determined the sample size necessary to detect a difference in alcohol-related youth injuries and deaths as a result of the intervention.

Quantity

Logistics are the methods, timing, and physical infra-structure for gathering and handling evidence. People and organizations have cultural preferences that dictate acceptable ways of asking questions and collecting infor-mation, and influence who is perceived as an appropriate person to ask the questions (i.e., someone known within the community versus a stranger from a local health agency). The techniques used to gather evidence in an evaluation must be in keeping with a given community's cultural norms. Data collection procedures should also protect confidentiality.

> PRAID coalition members determined that the study population of young men would be more likely to participate in focus groups if peers and youth leaders were used to market and facilitate the groups. The focus group meetings were scheduled in locations where young men in the town normally congregated.

Logistics

Analysis and synthesis are methods to discover and summarize evaluation findings. In evaluations that use multiple methods, patterns in evidence are detected by:
- isolating important findings (**analysis**)
- combining different sources of information to reach a larger understanding (**synthesis**)

Analysis and synthesis involve deciding how to organize, classify, compare, and display information. These decisions are guided by the questions being asked, the types of data available and, especially, by input from stakeholders and primary intended users.

> Members of the evaluation team examined pre- and post-intervention questionnaire data as well as qualitative interview data for analysis and synthesis. The following questions were raised by coalition members during this phase: What are the common

Analysis and Synthesis

ideas, themes, and patterns within the qualitative and quantitative data? What do the frequency data show? Are there any surprising findings? Did all available males within the three designated high schools participate in the education programs? The answers to these questions were then displayed graphically in charts that tracked changes in beliefs, behavior, and participation in alcohol education. Graphs were also created to show the number of citations issued to young men within each geographic area over a given period of time.

Applying standards

As we have noted in previous steps, you can help ensure that your evaluation is a quality one by considering relevant evaluation standards throughout the evaluation process. The two standards that most directly apply to Step 4—Gather and Analyze Evidence—are **utility** and **accuracy**. As you participate in gathering and analyzing evidence, the questions presented in the following table can help you to clarify and achieve these standards.

Standards for Step 4: Gather and Analyze Evidence

| Questions	Standards
◆ Have you reached an agreement on techniques to analyze and synthesize findings before data collection begins? ◆ Will the information collected address pertinent issues about the program and is this responsive to needs of your stakeholders?	**Utility:** Ensures that the evaluation is useful and answers questions that are directly relevant to users.

Standards for Step 4: Gather and Analyze Evidence

(Table continued from page 46)

Questions	Standards
◆ Are the sources of information used in the program described in adequate detail? ◆ Do the information-gathering procedures address internal validity and reliability issues? ◆ Is there a system in place for identifying and correcting errors? ◆ Has the process of analyzing quantitative and qualitative data been effective in answering your key evaluation questions?	**Accuracy:** Ensures that the evaluation findings are considered correct.

In summary, Step 4—Gather and Analyze Evidence—represents a process through which information about the program in which you are engaged can be gathered and synthesized for subsequent presentation. It ensures that the benefits of evaluation (the uses of this information) are clear to all stakeholders and that the processes followed meet everyone's agreement.

Think about a program evaluation in which you are involved. The following questions may help you think about how to approach Step 4—Gather and Analyze Evidence.
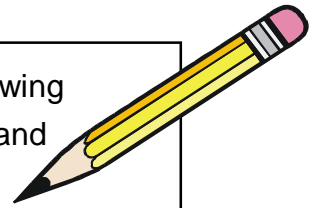
**What expertise and resources could you draw on for help in defining your methods?**

What sources of information could you use in the evaluation (people, documents, observations)?

What systematic processes could you use to gather information?

How could you check for errors and make corrections as part of the data gathering process?

What data analysis and synthesis techniques could you consider?

How could your analysis process answer key stakeholder questions effectively?

STEP 5

Justify Conclusions

STANDARDS

Utility   **Accuracy**
Propriety   **Feasibility**
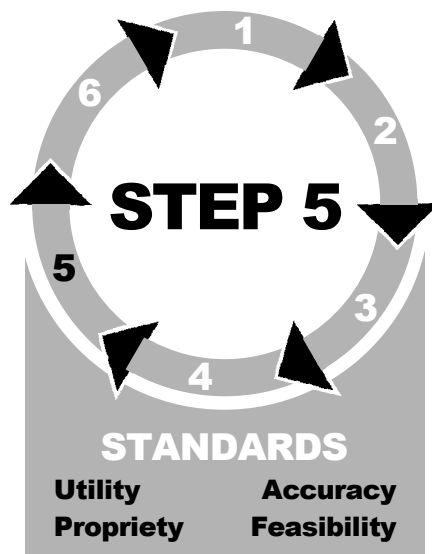
Definition

Justifying conclusions involves:

◆ making claims about the program based on the evidence gathered

◆ justifying the claims by comparing the evidence against stakeholder values

Why is it important to justify conclusions?

The evidence gathered for an evaluation does not necessarily speak for itself. To substantiate and justify conclusions, it is important to carefully consider the evidence from a variety of stakeholder perspectives. Conclusions become justified when they are linked to the gathered evidence and judged against values that stakeholders agree upon. When communities, agencies and other stakeholders agree that the conclusions are justified, they will use the evaluation results with more confidence.

Conclusions can be strengthened by searching for alternative explanations and showing why the evidence does not support them. Where there are different but equally well-supported conclusions, it can be helpful to present each set of conclusions with a summary of its strengths and weaknesses.

49

What's involved in justifying conclusions?

Four principal elements are involved in justifying conclusions:

| Program Standards | — | Interpretation | — | Judgments | — | Recommendations |
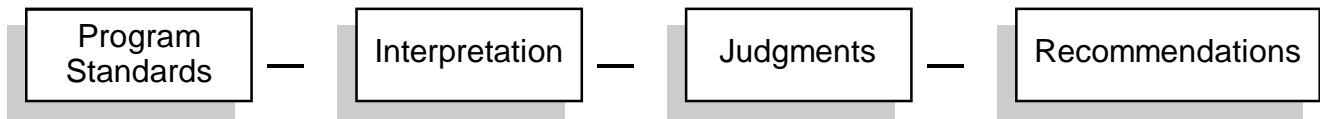
Program standards (not to be confused with the *evaluation* standards discussed throughout this document) are fundamental to sound evaluation. They reflect stake-holders' values about the program. Program standards:

◆ allow evaluators and stakeholders to make judgments about the program

◆ provide alternative ways to compare program results

When stakeholders articulate and negotiate their values, these become the standards for considering a program "successful," "adequate," or "unsuccessful." Examples of program standards that can be used to assess program results include community values or norms, professional standards, program objectives, fixed performance criteria, performance by previous or similar programs, considerations of feasibility and sustainability, social equity, and other stakeholder values.

> The central objective of the community-based intervention program was to reduce by 50% the incidence of alcohol-related motor vehicle crashes (MVCs) in males under 21 by the year 2005. Because an important purpose of the evaluation was to determine the program's effects, it made sense to the PRAID coalition to use this program objective as one of the standards for making judgments about the program. The objective of building coalitions and community partnerships represented another key means of judging the program.

Program Standards

Interpretation is the effort to figure out what the evaluation findings mean. Interpretation:
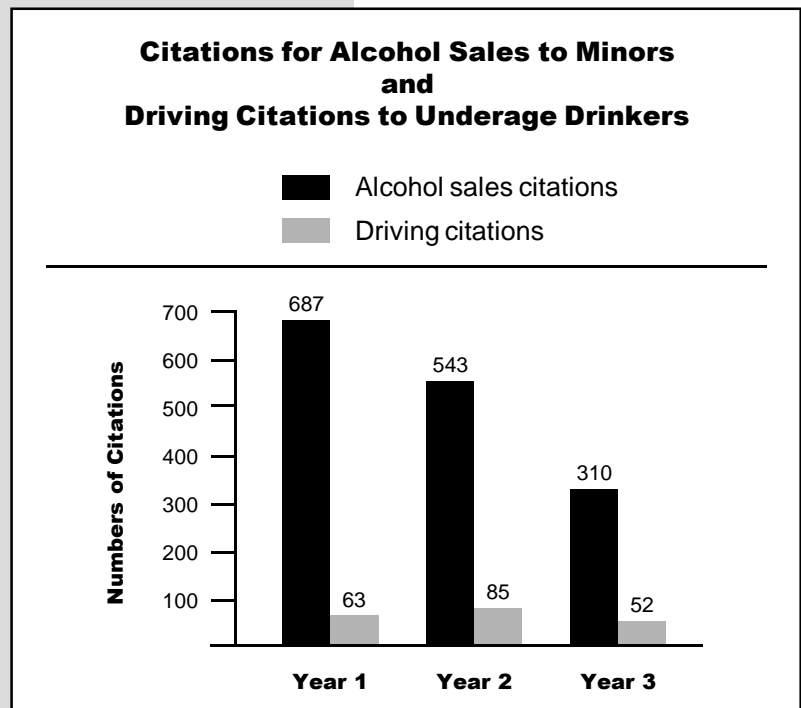
- ◆ draws on information and perspectives that stakeholders bring to the evaluation
- ◆ is strengthened through active stakeholder participation and interaction

Uncovering facts about a program's performance is not enough to draw conclusions. The facts must be interpreted to understand their practical significance.

Interpretation

Imagine that we have stepped three years into the future. At that point in time, the evaluation findings show a modest reduction in alcohol-related motor vehicle crashes among male youth in the project's geographic area. As the PRAID coalition gathers to interpret these findings, they may try to figure out whether each of the three broad program strategies (increased enforcement/sanctions, comprehensive education, reduced alcohol sales) has had an impact on the documented decline in alcohol-related crashes. Alternatively, the coalition members may focus on interpreting the combined contributions of the set of intervention strategies, rather than zeroing in on any single strategy.

The graph illustrates data collected over a three-year period. It summarizes citations issued to alcohol vendors (e.g., liquor stores, bars) for sales of alcohol to minors and alcohol-related driving citations issued to underage drinkers in the

Citations for Alcohol Sales to Minors and Driving Citations to Underage Drinkers

- ■ Alcohol sales citations
- ▨ Driving citations

	Year 1	Year 2	Year 3
Alcohol sales citations	687	543	310
Driving citations	63	85	52

Numbers of Citations: 700, 600, 500, 400, 300, 200, 100

City of Hope (population 22,781). Data can be analyzed and interpreted to ascertain whether collective intervention strategies have had any effect on alcohol-related MVCs.

Judgments are statements about a program's merit, worth, or significance. They are formed when findings and interpretations are compared against one or more selected program standards. In forming judgments about a program:

- multiple program standards can be applied
- stakeholders may reach different or even conflicting judgments

Conflicting claims about a program's quality, value, or importance often indicate that stakeholders are using different program standards or values in making their judgments. This type of disagreement can prompt stakeholders to clarify their values and reach consensus on how the program should be judged.

The evaluation documented an increase in highway patrols for underage drinking and driving in some areas of the county on selected weekends. Police department stakeholders, who viewed any improvement in performance as a sign of success, viewed this increase as a reason to judge the program positively. However, coalition members from MADD, SADD, and community-based organizations had an expectation that countywide patrols would be increased every weekend; as a result, their judgment of program performance was less positive than that of the police.

Judgments

52

Recommendations are actions to consider as a result of an evaluation. Recommendations require information beyond what is necessary to form program judgments. Recommendations:

- ◆ can strengthen an evaluation when they anticipate and react to what users want to know
- ◆ may undermine an evaluation's credibility if they are not supported by enough evidence, or are not in keeping with stakeholders' values

The chances that recommendations will be relevant and well received can be increased by sharing draft recommendations, soliciting reactions from multiple stakeholders, and presenting options instead of directive advice.

> After holding a community forum to further interpret the evaluation results and make judgments about the program, the PRAID coalition members used the results to come up with recommendations that they presented to the county commissioners and school boards. For example, because the evaluation suggested that police patrols on weekends had had an impact on young men's drinking and driving behavior, the coalition recommended that the level of police patrolling be permanently increased. However, the evaluation results showed low levels of participation by high school students in the program's educational initiatives, leading the coalition members to recommend that community-based (rather than school-based) strategies be developed to raise young men's awareness about the dangers of alcohol.

Recommendations

Applying standards

As we have noted in the previous steps, you can help ensure that your evaluation is a quality one by considering relevant evaluation standards throughout the evaluation process. The two standards that most directly apply to

Step 5—Justify Conclusions—are **utility** and **accuracy**. As you participate in justifying conclusions, the questions presented in the following table can help you to clarify and achieve these standards.

Standards for Step 5: Justify Conclusions

Questions	Standards
◆ Have you carefully described the perspectives, procedures and rationale used to interpret the findings? ◆ Have stakeholders considered different approaches for interpreting the findings?	**Utility:** Ensures that the evaluation is useful and answers questions that are directly relevant to users.
◆ Can you explicitly justify your conclusions? ◆ Are the conclusions fully understandable to stakeholders?	**Accuracy:** Ensures that the evaluation findings are considered correct.

In summary, Step 5—Justify Conclusions—involves interpreting evaluation results so that they make sense to all stakeholders and reflect stakeholders' values about what is important. Completing this step can help stakeholders consider what actions to take as a result of the evaluation.

Think about a program evaluation in which you are involved. The following questions may help you approach how you think about Step 5—Justify Conclusions.
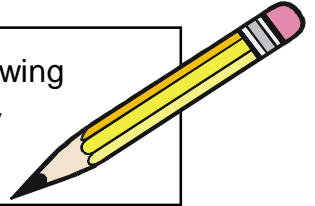
How could you involve stakeholders in interpreting evaluation results for your program?

If stakeholders had conflicting judgments about your program, how could you reach consensus?

What information could you use to develop recommendations for action?

How might you share recommendations with other stakeholders?

STEP 6

Ensure Use and Share Lessons Learned

STANDARDS

Utility   Accuracy
Propriety   Feasibility

Definition

Using evaluation results appropriately means:

◆ making stakeholders aware of the evaluation procedures and findings

◆ considering findings in program-related decisions or actions

◆ ensuring that those who participate in the evaluation perceive it as beneficial

Why is it important to ensure use and share lessons learned?

It is naive to assume that lessons learned in an evaluation will necessarily be used in decision-making and subsequent action. Evaluation participants must make a deliberate effort to promote use of the evaluation findings. They also have a responsibility to prevent misuse of findings. Factors that influence whether results are used appropriately include: evaluator credibility; disclosure of findings; report clarity, timeliness, and impartiality; and changes in the program or organization context. Ensuring use requires thinking strategically from the earliest stages of the evaluation, as well as diligently looking for opportunities to communicate results and influence program decisions and/or policy makers.

[Evaluation] is a good thing to do for a lot of reasons. But let us be clear that many decisions get made politically and have little or nothing to do with good data and the value of what's going on.
CBO Director

The very process of doing evaluation is important. When individuals are exposed to the logic, reasoning, and values that guide evaluation, their thinking and behavior may change profoundly. Participation in an evaluation may:

- encourage stakeholders to base decisions on systematic judgments instead of on unfounded assumptions
- prompt stakeholders to clarify their understanding of program goals, thereby improving their ability to function as a team
- help stakeholders clarify what is really important through the process of defining indicators
- make outcomes matter by changing the reinforcements connected with achieving positive results, as when funders offer "bonus grants" or "outcome dividends" to programs that show significant community change and improvement

What's involved in ensuring use and sharing lessons learned?

Four elements are important in making sure that the findings from an evaluation are used:

| Preparation | — | Feedback | — | Follow-up | — | Dissemination |
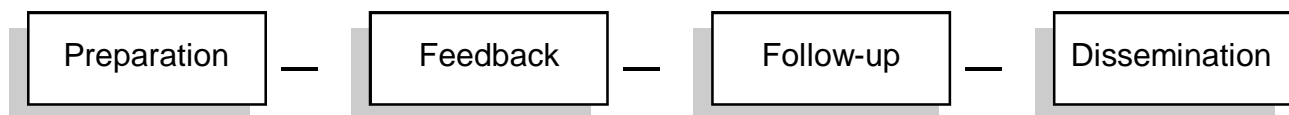
Preparation refers to the steps taken to get ready to eventually use the evaluation findings. Through preparation, stakeholders can:

- strengthen their ability to translate new knowledge into appropriate action
- discuss how potential findings might affect decision-making
- explore positive and negative implications of potential results and identify different options for program improvement

The preliminary evidence of increased highway patrols for underage drinking and driving drew mixed reactions from various members of the PRAID coalition. Police department representatives perceived the staffing levels required for increased patrolling to be unsustainable due to competing priorities and public safety requirements, whereas citizen and MADD representatives thought the benefits of the patrols were probably worth the tradeoffs. By preparing for the evaluation results early on, the primary users and other stakeholders were able to begin discussing possible decisions and actions.

Feedback is the communication that occurs among everyone involved in the evaluation. Feedback, necessary at all stages of the evaluation process, creates an atmosphere of trust among stakeholders. Early in an evaluation, the process of giving and receiving feedback keeps an evaluation on track by keeping everyone informed about how the program is being implemented and how the evaluation is proceeding. As the evaluation progresses and preliminary results become available, feedback helps ensure that primary intended users and other stakeholders have opportunities to comment on evaluation decisions. Valuable feedback can be obtained from stakeholders by holding discussions during each step of the evaluation and routinely sharing interim findings, provisional interpretations, and draft reports.

To provide feedback to and obtain feedback from each type of stakeholder involved in the PRAID coalition, an evaluation task force was formed, coordinated by a funded staff person. From the earliest stages of the evaluation, the staff person planned and coordinated quarterly meetings for task force members, who then went back to their own groups (e.g., high schools, police, the local health department,

Feedback

58

hospitals, beverage distributors, MADD, SADD, other community-based organizations) to share program progress. As the evaluation progressed and preliminary evaluation results became available, the staff person also used a variety of other approaches to ensure that all the coalition members remained well informed and able to provide feedback. These included more frequent updates through e-mail, Web-based communications, and interim print reports, as well as meetings and discussions with the different groups represented by the task force.

Although **follow-up** refers to the support that many users need *throughout* the evaluation process, this step, in particular, refers to the support that is needed *after* users receive evaluation results and begin to reach and justify their conclusions. Active follow-up:

- ◆ reminds users of the intended uses of what has been learned
- ◆ can help to prevent misuse of results by ensuring that evidence is applied to the questions that were the evaluation's central focus
- ◆ prevents lessons learned from becoming lost or ignored in the process of making complex or political decisions

As evaluation results for the intervention were compiled and analyzed, the task force staff person continually reminded task force members of the evaluation's intended uses. Although health care providers and citizens were interested in considering the implications of results showing a decline in alcohol sales to underage males, other coalition members such as liquor distributors and city officials (concerned with alcohol-related tax revenues) sought to draw attention away from such findings. The follow-up of the staff person helped keep discussion focused on the most important evaluation results.

Follow-up

Dissemination is the process of communicating evaluation procedures or lessons learned to relevant audiences in a timely, unbiased, and consistent manner. Regardless of how communications are structured, the goal for dissemination is to achieve full disclosure and impartial reporting. Planning effective communications requires:

- advance discussion of the reporting strategy with intended users and other stakeholders
- consideration of the timing, style, tone, message source, vehicle, and format of information products

> The project produced and disseminated evaluation reports that had limited technical jargon and used lots of graphics, stories, and examples. In addition to using reports and other written products for dissemination, PRAID coalition members sustained momentum and broadened the evaluation's reach by making presentations at school board meetings, gatherings of public officials, and other community events. Community members used the project's lessons to begin planning and developing new community-based initiatives.

Dissemination

I left a meeting on the evaluation of Project X and it wasn't for the community at all. It was really for policy makers. . . . If you really want people to use it, [you need to pay attention to] the use of language and using other multiple media formats for this real complex information . . . .

CBO Director

Applying standards

As we have already noted, you can help ensure the quality of your evaluation by considering relevant evaluation standards throughout the evaluation process. The three standards that most directly apply to Step 6—Ensure Use and Share Lessons Learned—are **utility**, **propriety**, and **accuracy**. As you use your own evaluation results, the questions presented in the following table can help you to clarify and achieve these standards.

Standards for Step 6: Ensure Use and Share Lessons Learned

Questions	Standards
◆ Are participants in the evaluation group trustworthy and competent? ◆ Do reports clearly describe the program, including its context, and the evaluation's purposes, procedures, and findings? ◆ Have you shared significant mid-course findings and reports with users so that the findings can be used in a timely fashion? ◆ Have you planned, conducted, and reported the evaluation in ways that encourage follow-through by stakeholders?	**Utility:** Ensures that the evaluation is useful and answers questions that are directly relevant to users.
◆ Have you ensured that the evaluation findings (including the limitations) are made accessible to everyone affected by the evaluation and others who have the right to receive the results?	**Propriety:** Ensures that the evaluation is ethical, conducted with regard for the rights and interests of those involved.
◆ Do the evaluation reports impartially and fairly reflect evaluation findings? ◆ Have you tried to avoid the distortions that can be caused by personal feelings and other biases?	**Accuracy:** Ensures that the evaluation findings are considered correct.

In summary, Step 6—Ensure Use and Share Lessons Learned—emphasizes the importance of translating results into action. Through the process of preparation, feedback, follow-up, and dissemination, the benefits of evaluation become clear to stakeholders. This step ensures that the lessons learned from an evaluation are shared in such a way as to influence program decisions, policy makers, and community-based initiatives.

Think about a program evaluation in which you are involved. The following questions may help you think about how to approach Step 6—Ensure Use and Share Lessons Learned.
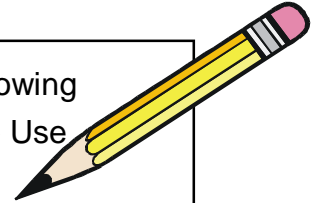
How could you ensure that stakeholders receive and provide feedback throughout the evaluation process?

How could you make sure that lessons learned are used?

What support is available to follow up on evaluation results?

What types of communication strategies might be appropriate for your program and stakeholders?

Glossary

The following glossary of terms is a compilation of definitions from numerous evaluation sources.

Community Capacity
The commitment, resources, and skills that a community can mobilize and use to address community issues and problems and strengthen community assets; the characteristics of communities that affect their ability to identify and address social and economic health issues; the cultivation and use of transferable knowledge, skills, systems, and other resources to affect community—and individual—level change.

Data
Information produced by and used in evaluation—includes numbers, words, pictures, or any type of information used.

Evaluation Design
A blueprint, strategy, or outline to answer questions about a program. Includes a clear statement about the purpose and plans for gathering, processing, and interpreting the information needed.

Evaluation Methods
Data collection options and strategies selected to match or fit the overall design and answer the evaluation questions. Methods depend on knowing who the information is for, how it will be used, what types of information are needed and when, and the resources available.

Formative Evaluation
A subset of process evaluation. Information collected for a specific period of time, often during the start-up or pilot phase of a project, to refine and improve implementation and solve unanticipated problems. This can include program monitoring efforts to provide ongoing feedback and assess intermediate outcomes.

Impact Evaluation	Assesses whether a program has achieved desired intermediate changes in individuals, population groups, or organizations.
Indicators	Benchmarks used to measure or test changes. Indicators can be at the level of individuals (e.g., behavioral changes), organizations (e.g., service delivery hours), or communities (e.g., unemployment rates, quality of life variables).
Outcome Evaluation	Examines the effects of a program on health status, usually defined in terms of morbidity (illness, injury) and mortality (death) rates. Determines the long-term effects of a program or intervention.
Process Evaluation	Addresses questions related to how a program is implemented. Compares what was supposed to happen with what actually happened. Answers questions about why the program succeeded, failed, or requires revising.
Program	A series of activities supported by a group of resources intended to achieve specific outcomes among particular groups.
Program Evaluation	The systematic collection, analysis, and reporting of information about the activities, characteristics, and outcomes of programs to make judgements about the program, improve program effectiveness, and/or inform decisions about future programming. Utilization-focused program evaluation is done for and with specific, intended primary users for specific, intended uses.
Qualitative Data	Information gathered from interviews, observations, or documents. May include detailed descriptions of situations, events, people, interactions, observed behaviors, and people's own thoughts about their experiences, attitudes, and beliefs. Other data sources are excerpts or entire passages from documents, correspondence, records and case histories.

Quantitative Data Information from questionnaires, tests, standardized (fixed, unchanging) observation instruments, and program records. Focuses on things that can be counted, categorized, and subjected to statistical analysis.

Sample Size A subset of cases (e.g., individuals, records, communities) selected from a population. In quantitative research, large samples generally enhance confidence in survey results. A minimum total number in a sample is important for some statistical tests.

Stakeholders People who care about what will be learned from the evaluation and about what will be done with the knowledge gained.

Summative Evaluation Conducted after program completion, often for the benefit of external, decision-makers (e.g., funding agencies, oversight offices, other users). Provides data to support judgements about a program's worth so that decisions can be made about continuation or expansion.

Themes Patterns or recurrent ideas that emerge as qualitative data are collected, analyzed, and interpreted.

Units of Analysis The primary focus of data collection and analysis. Can be individuals, groups of people (e.g., classrooms), neighborhoods, cities, states, or countries. Units may also be events or incidents (e.g., service delivery failures). Each unit may require different kinds of data. The focus of analysis also varies and affects what can be said (e.g., if the unit of analysis is a city, one cannot make a statement about the impact of the program on an individual).

Program Evaluation Resources

Web-based resources:

Centers for Disease Control and Prevention: http://www.cdc.gov/eval/

W.K. Kellogg Foundation: http://www.wkkf.org/Publications/evalhdbk/

University of Kansas: http://ctb.lsi.ukans.edu/ctb

University of Toronto: http://www.utoronto.ca/shp/hcu

University of Wisconsin Extension: http://www.uwex.edu/ces/pdante/evaluat.htm/

Publications

Connell JP, Kubisch AC, Schorr LB, Weiss, CH. *New Approaches to Evaluating Community Initiatives,* New York, NY: Aspen Institute, 1995.

Fawcett SB, Paine-Andrews A, Francisco VT, Schulz J, Ritchter KP, et al. Evaluating community initiatives for health and development. In *Evaluating Health Promotion Approaches,* edited by I Rootman and D McQueen. Copenhagen, Denmark: World Health Organization — in press.

Fawcett SB, Sterling TD, Paine Andrews A, Harris KJ, Francisco VT, et al. *Evaluating Community Efforts to Prevent Cardiovascular Diseases.* Atlanta, GA, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 1995.

Fetterman DM, Kaftarian SJ, Wandersman A. *Empowerment Evaluation: Knowledge and Tools for Self-assessment and Accountability.* Thousand Oaks, CA: Sage Publications, 1996.

Patton MQ, *Utilization-focused Evaluation.* Thousand Oaks, CA: Sage Publications, 1997.

Rossi PH, Freeman HE, Lipsey MW. *Evaluation: A Systematic Approach.* Newbury Park, CA: Sage Publications, 1999.

Shadish WR, Cook TD, Leviton LC. *Foundations of Program Evaluation.* Newbury Park, CA: Sage Publications, 1991.

Taylor-Powell E, Steele S, Douglas M. *Planning a program evaluation.* Madison, WI: University of Wisconsin Cooperative Extension, 1996 (see web based entry on page 66).

University of Toronto, Health Communication Unit at the Center for Health Promotion, *Evaluating Health Promotion Programs* (see web based entry on page 66).