
Evaluating HIV/AIDS Prevention Projects

A Manual for Nongovernmental Organizations

Jane T. Bertrand, Ph.D., MBA

MEASURE Evaluation

Manuel Solís, M.D.

Guatemalan Association for AIDS
Prevention and Control (AGPCS)

The manual series is made possible by support from USAID under the terms of Cooperative Agreement HRN-A-00-97-00018-00. The opinions expressed are those of the authors, and do not necessarily reflect the views of USAID.

Recommended citation:

Bertrand, Jane T., and Manuel Solís. Evaluating HIV/AIDS Prevention Projects: A Manual for Nongovernmental Organizations. **MEASURE *Evaluation Manual Series*, No. 10**. Carolina Population Center, University of North Carolina at Chapel Hill. July 2000. (Translation to English 2004).

Acknowledgments

MEASURE *Evaluation* and the Guatemalan Association for AIDS Prevention and Control (AGPCS, subsequently renamed ASI, the Association for Integrated Health) produced this manual under the United States Agency for International Development (USAID) cooperative agreement HRN-A-00-97-00018-00. Special thanks to Stanley Terrell, USAID/G-CAP, who supported this work from the beginning.

The authors originally published this manual in Spanish under the title *Evaluando Proyectos de Prevención de VIH/SIDA: Un Manual con Enfoque en las Organizaciones No Gubernamentales*, with a particular focus on Central America. The authors have adapted it to be applicable to HIV/AIDS programs worldwide. We mention nongovernmental organizations (NGOs) in the title, because NGOs often target activities and services to geographically limited areas or specialized subgroups of the population. Their interventions often do not lend themselves to national-level population-based evaluation. In fact, this manual is intended for any organization — public or private — that implements projects rather than national-level programs. For large-scale HIV/AIDS programs, see *National AIDS Programmes: A Guide to Monitoring and Evaluation*, (UNAIDS et al., 2000).

The authors wish to thank several collaborators working in HIV/AIDS in Central America at the time we wrote the original version of this manual: Annie de Salazar (Director of ASI, Guatemala), Ann Fitzgerald and Eugenia Monterroso (Directors from 1999–2000 and 2000–present, respectively, of the Central

American AIDS Prevention Program, Proyecto Accion SIDA de Centro America [PASCA]). We also thank Jeannette de Jiménez of PASCA for her collaboration on this manual.

From MEASURE *Evaluation*, Gabriela Escudero, Sandra Bedoya, and Verónica Arroyave assisted in editing the manual. They, as well as Stephen Sosler and Ebony Davis, participated in the development of Chapter 5 describing different illustrative HIV/AIDS programs, their indicators, and appropriate data sources. Thanks also to Marlene Vera who helped in editing the Spanish version of the document, and Dr. Robert Magnani for his review of Appendix B on sampling.

Several colleagues reviewed earlier drafts of the Spanish version of this manual and offered valuable suggestions on the content, including: Mario Bronfman, Rafael Mazín, Annie de Salazar, Ingrid Sanchinelli, Edward Scholl, and Stanley Terrell. Kristin Kaper prepared the first draft of this English translation; Kathy Wolfe assisted in manuscript preparation; Donna Espeut and Kavita Singh made modifications to the English translation to make it applicable to NGOs worldwide; and the lead author, Jane Bertrand, had oversight for finalizing the document.

The content in the manual remains very similar to the original Spanish version published in 2000. The translation is not literal, but the authors have remained faithful to the concepts in the original version. The English version includes additional indicators to measure the A and B of the ABCs (abstinence, be faithful, and condom use) and an additional illustrative intervention related to the prevention of mother-to-child transmission (PMTCT) of HIV/AIDS.

Table of Contents

1	Evaluating HIV/AIDS Prevention Programs	9
	What Is Evaluation?	9
	What Is the Purpose of Evaluation?	9
	Why Do People Avoid Evaluation?	10
	The Purpose of This Manual	11
2	Concepts and Definitions	13
	Combating AIDS	13
	Programs, Projects, Interventions, and Activities	15
	Levels of Evaluation	18
	Evaluation at the Population Level	18
	Evaluation at the Program Level	19
	What Do We Want to Learn from an Evaluation?	20
	Process Evaluation	20
	Monitoring of Results (Outputs, Outcomes)	21
	Impact Assessment (Cause and Effect)	23
	Example of the Three Types of Evaluation Applied to an Illustrative NGO Project	24
	Example 1: Process Evaluation	24
	Example 2: Monitoring of Results (Change)	25
	Example 3: Impact Assessment (To Demonstrate Cause and Effect)	27
	Types of Evaluations Conducted to Date on HIV/AIDS Prevention Projects	29

3	Designing an Evaluation	31
	Steps in Designing an Evaluation	31
	Step 1: Determine (or Clarify) the Project’s Objectives	32
	Step 2: Determine the Type of Evaluation to Be Carried Out	34
	Step 3: Identify Indicators for Each Type of Evaluation	35
	Step 4: Determine the Appropriate Data Source for Each Indicator	38
	Step 5: Prepare an Evaluation Plan	39
	Some Suggestions for Evaluating Projects with Multiple Components	41
	Option 1: Presentation by Subgroup (Different Intended Audiences)	41
	Option 2: Presentation by Methodology Used	42
	Option 3: Presentation by Type of Intervention	43
4	Methodologies for Collecting Data	45
	What Is the Difference Between Quantitative and Qualitative Methodologies?	45
	Quantitative Evaluation Methods	46
	Structured Interviews	46
	Service Statistics	50
	Qualitative Methods	52
	Observation	52
	Direct Observation (with a Guide or Checklist)	52
	Systematic Observation by a Mystery (or Simulated) Client	55
	Ethnographic Observation	57
	In-Depth Interviews	58
	Focus Groups	60

Example of the Use of Qualitative and Quantitative Methods: Evaluating Client Satisfaction	64
Methodological Note about Focus Groups	70
Linking the Type of Evaluation to the Appropriate Data Source	71
5 HIV/AIDS Illustrative Projects	73
1: Telephone Hotline on STI/HIV/AIDS	76
2: School-based Education on STIs and HIV/AIDS for Youth	78
3: Peer Education	80
4: National or Regional HIV/AIDS Information Centers	82
5: Voluntary Counseling and Testing (VCT)	84
6: Integration of STI/HIV/AIDS into MCH and Family Planning Programs	86
7: Prevention of Mother-to-Child Transmission (PMTCT) of HIV	88
8: Condom Social Marketing	90
9: Empowerment of Female Commercial Sex Workers (CSWs)	92
10: Promotion of Condom Use among High-risk (and Often Predominantly Male) Audiences (Clients of CSWs, Men Who Have Sex with Men, Prisoners, Military Personnel)	94

6 Data Processing and Analysis	97
Quantitative Data	97
Editing	97
Coding	98
Data Source	99
Creating a Database	99
Data Entry	99
Data Cleaning	100
Data Analysis	100
Illustrative Presentation of Quantitative Results	103
Qualitative Data	112
Preparing the Transcriptions	112
Processing Data for Analysis	113
Preparing the Report	114
Conclusions and Recommendations	115
Using Evaluation Results to Improve Programs	115
Audiences for Evaluation Results	115
Some Formats for the Presentation of Results	117
 Appendix A: An Inventory of Common Indicators	 121
Appendix B: Notes on Sampling	128
Appendix C: Additional Bibliographic References	144

Chapter 1

Evaluating HIV/AIDS Prevention Programs

Most HIV/AIDS programs aim to prevent the transmission of the HIV virus and to improve the quality of life for those who are infected or affected by HIV/AIDS.

What Is Evaluation?

Evaluation is the systematic application of quantitative and qualitative research techniques to determine the appropriateness and effectiveness of the design and implementation of social programs. Often in implementing programs, we become so caught up in the day to day challenges that we neglect asking a key question: “Is our program making a difference?” More specifically, “Is it achieving the stated objectives?” Evaluation offers the answer to this question.

What Is the Purpose of Evaluation?

Evaluation provides many benefits to social programs. HIV/AIDS prevention is no exception. Evaluation permits us to:

- identify successful strategies;
- modify or discontinue interventions that do not yield desired results;

-
-
- share findings with programs in other countries or regions;
 - provide donors with evidence of the results of their investment; and
 - demonstrate the organization's interest in accountability.

Evaluation helps program managers identify what is and is not working, as well as how to make the project work better. Evaluation also provides a means of demonstrating to project staff and donor agencies the extent to which a project is achieving its objectives.

Why Do People Avoid Evaluation?

Often project staff are so busy confronting the challenges of running an HIV/AIDS project that they overlook evaluation or put it off until "later." Others may not consider it a priority, and many are apprehensive of evaluation for different reasons:

- They are convinced that they are doing a good job, so think, "Why bother to evaluate?"
- They believe evaluation is tantamount to "inspection."
- They fear negative feedback. They are so committed to what they are doing that they don't want to consider that it might not be producing the desired results.
- They prefer to devote their time and energy to new programmatic activities, not to studying past results.
- They don't have the financial or material resources to conduct evaluation.
- They don't have personnel with training or experience in evaluation.

The Purpose of This Manual

The purpose of this manual is to demystify the evaluation process, especially for staff who are not specialized in evaluation techniques. It is intended to serve as a guide for designing and implementing evaluations of HIV/AIDS prevention projects.

Our principal audience for the manual is nongovernmental organizations, or NGOs. The distinction is not so much the type of organization in question (e.g., NGO), but rather the expected coverage and nature of their programs. Often NGOs work with specific subgroups within the population (e.g., adolescents, commercial sex workers, men who have sex with men) or in a limited geographical area (e.g., the capital city, a remote set of rural villages, or a specific region or district). They also devote a substantial amount of attention to addressing sociocultural factors and processes that can impact behavior change, and they usually employ community empowerment approaches to solve local problems.

Although exceptions exist, NGOs generally do not aspire to have measurable effects at the national level. Another manual entitled *National AIDS Programs: A Guide to Monitoring and Evaluation*, produced by UNAIDS in collaboration with MEASURE *Evaluation* and other partners (2000), provides an excellent set of indicators for programs at the national level. By contrast, the current manual is designed for organizations whose programs and projects are smaller in scope, most notably NGOs.

The majority of NGOs that work in AIDS prevention share a common objective: to slow — and ultimately stop — the spread of the AIDS epidemic in countries worldwide. To meet this challenge, we must find effective strategies for reducing transmission and caring for those infected and affected by HIV/

AIDS. Evaluation takes on special importance if we consider it the number one means of improving on what we do in search of this common objective.

Chapter 2

Concepts and Definitions

This chapter provides a general overview of the concepts and terminology frequently used in monitoring and evaluating projects. It attempts to explain these concepts in a language that nonspecialists will find useful.

Combating AIDS

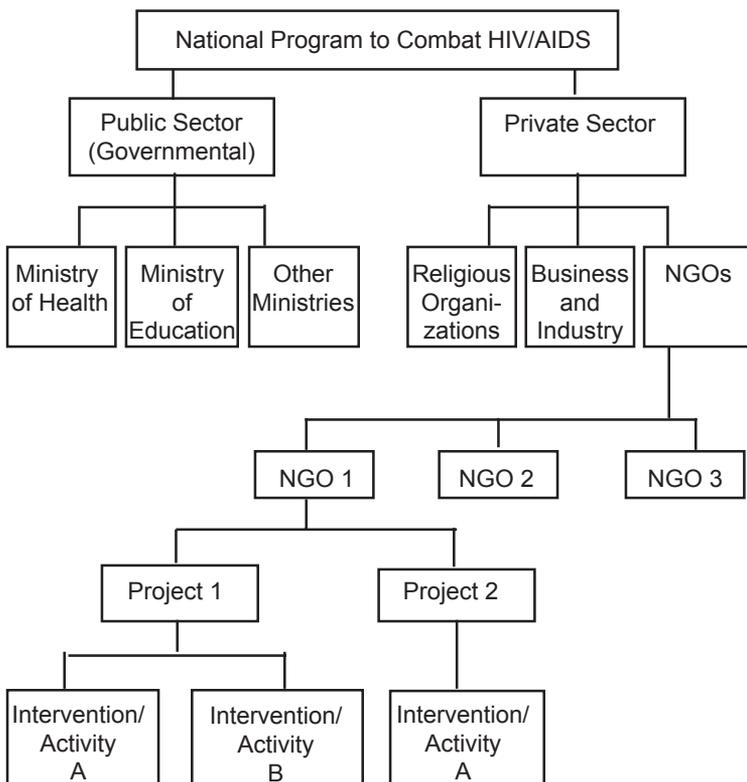
The main objective of HIV/AIDS prevention programs is to halt the spread of HIV infection. Given that neither a vaccine nor cure exists, this ambitious goal will require the sustained efforts of many groups for many more years.

In most developing countries, both the government and private sector participate in the fight against AIDS. The governmental sector consists of different ministries (e.g., health, education, information, youth, women's affairs, transport). The private sector comprises businesses and industries, commercial or social marketing programs, and a range of nongovernmental organizations (NGOs).

In some countries, a formal coalition exists to coordinate the work of various organizations involved in HIV/AIDS, including the government. This group of organizations constitutes the National Program for HIV/AIDS. In other countries, no formal coalition exists; rather, one or more organizations work in a parallel or isolated manner in prevention, counseling and testing, treatment, and care and

support. Whatever the configuration, all of the organizations (public and private) that work in the field of HIV/AIDS within a country collectively constitute the national program. Figure 1 illustrates both public and private sectors and their respective institutions that may work in HIV/AIDS.

Figure 1. Illustrative Structure of a National HIV/AIDS Program



Programs, Projects, Interventions, and Activities

Each organization that forms part of the national program carries out one or more HIV/AIDS prevention, counseling and testing, or care/support strategies, which may take the form of a program, project, or activity. The definitions of these terms vary among countries and institutions. The definitions used for these terms in this manual appear in *Boxes 1, 2, and 3*.

It is helpful to recognize the hierarchy among these terms. For example, a given NGO may have an HIV **program**, which consists of three separate **projects**, each funded by a different donor (e.g., a school-based project to impart life skills education, a community-based condom distribution project, and a voluntary counseling and testing service). “Program” is the broadest term, since program can include a number of different projects, such as training of trainers, curriculum development, preparation of audiovisual aids, and realization of the workshops, each financed by a different donor. In turn, each project has at least one intervention or strategy to reach the target population.

In this manual, we tend to focus on **projects**, which contribute to the overall goal of reducing HIV/AIDS and improving the quality of life for those infected or affected. Yet projects — because of their limited scope and duration — tend not to have a national-level impact by themselves.

One purpose of evaluation is to determine if a given program or project has achieved its objectives. It is also important to understand the process of implementation, in order to identify successful aspects and to improve (or eliminate, if necessary) ineffective components.

Despite the distinction we have made between “program” and “projects” in the context of NGO work, we frequently use the term program evaluation in this manual. This term covers a wide range of research and evaluation activities (for example, all that are presented in this manual), and the principles presented are equally applicable to programs or projects.

Box 1: Definition of a Program

A combination of interventions or activities that an organization establishes as a fundamental part of its structure and mission. Programs tend to focus on a specific area (e.g., HIV/AIDS) and operate over the long-term. Organizations develop programs consistent with their mission and policies.

Box 2: Definition of a Project

A combination of interventions or activities that an organization establishes in response to specific circumstances or needs. Projects tend to have a defined duration (e.g., 3–5 years). If a project is of particular importance to an organization, it may evolve into a program. For example, an NGO that starts with a single HIV/AIDS activity may go on to develop a number of related HIV/AIDS projects, which evolve into the NGO’s program for HIV/AIDS. Box 4 lists illustrative HIV/AIDS projects that some NGOs implement.

Box 3: Definition of an Activity/Intervention

Specific actions or events implemented with the intention of reaching a given audience. Activities are what the program or project does for the purpose of achieving its objectives. A given activity — when frequently repeated — constitutes an intervention or strategy (e.g., mass media programming, in-school life skills education).

Box 4: Illustrative Projects for Combating HIV/AIDS

Communication, Information, Education/ Behavior Change	<ol style="list-style-type: none">1. Telephone Hotline on STI/HIV/AIDS2. School-based Education on STIs and HIV/AIDS for Youth3. Peer Education
Multi-Sectoral Services/Clinical	<ol style="list-style-type: none">4. National or Regional HIV/AIDS Information Centers5. Voluntary Counseling and Testing (VCT)6. Integration of STI/HIV/AIDS into MCH and Family Planning Programs7. Prevention of Mother-to-Child Transmission (PMTCT) of HIV
Social Marketing	<ol style="list-style-type: none">8. Condom Social Marketing
Projects for Vulnerable Populations	<ol style="list-style-type: none">9. Empowerment of Female Commercial Sex Workers (CSWs)10. Promotion of Condom Use Among High-risk (and Often Predominantly Male) Audiences (e.g., Clients of CSWs, Men Who Have Sex with Men, Prisoners, and Military Personnel)

The projects listed in Box 4 are illustrative of the types of projects that NGOs have been implementing for many years; NGOs are certainly not limited to the above interventions. For example, a number of NGOs are involved in social and behavioral change campaigns that use mass media (such as radio programming) and/or traditional media (e.g., street theatre, and storytelling) to address social norms, myths, and perceptions related to HIV/AIDS. Other NGOs, including faith-based organizations, engage community leaders to become facilitators of behavior change, promoting HIV risk-reduction and encouraging the compassionate treatment of individuals and families affected by HIV/AIDS. Some are involved in strengthening district-level responses to HIV/AIDS, while others are involved in national-level advocacy with other NGOs.

Some NGOs are beginning to implement projects focusing on universal precautions training; community-based distribution (CBD) programs to promote HIV prevention messages and increase access to dual methods (male/female condoms); and tailored HIV/AIDS prevention interventions that address issues which place orphans and vulnerable youth at risk. Community mobilization is also a key aspect of many projects. In Chapter 5, the 10 projects listed in Box 4 will help to illustrate the various concepts introduced in this manual.

Levels of Evaluation

An important first step in designing an evaluation is to decide if it will be performed at the population or program level. Key factors in the decision are the intended audience and the expected reach of the intervention.

Evaluation at the Population Level

This type of evaluation relates to the entire population of a given geographical area (e.g., city, district, country) who fit the

profile for the intended audience, **whether or not they participated in or were exposed to program activities**. Thus, it answers the question, “How effective was this intervention in reaching (and changing behavior) among the intended audience as a whole?” This type of evaluation is appropriate for large-scale programs designed to be far-reaching, such as behavior change programs that use the mass media in an attempt to reach the general public.

Evaluation at the Program Level

By contrast, NGO projects on HIV/AIDS are often smaller in scope, focusing on subgroups in the population with specific characteristics: adolescents in schools, military personnel, commercial sex workers, truck drivers, factory workers, seropositive individuals, and other definable groups in a determined geographical area. Program-level evaluation involves only those persons exposed to the program activities (e.g., participants in training courses, persons attending a VCT service, persons residing in target communities). This type of evaluation answers the question: “How effective was this intervention in changing behavior among those exposed to it?” Technically, one could use the term project-level evaluation with respect to projects, although “program-level evaluation” applies to both programs and projects.

NGOs often conduct program-level rather than population-level evaluation, because programs or projects targeted at a specific group in a defined geographical area may not have measurable effects at the population level. Moreover, population-based evaluation tends to be costly and is not justified unless the intervention is aimed at reaching a broad segment of the population.

What Do We Want to Learn from an Evaluation?

Evaluation can answer three basic questions:

1. How well has the project been implemented?
2. Has the desired change been achieved?
3. If the change has been achieved, to what extent can the change be attributed to the project?

Depending upon which question we want to answer, we choose one of three evaluation types:

1. process evaluation;
2. monitoring of results (outputs and outcomes); and
3. impact assessment (measuring cause-and-effect).

Another type of evaluation is cost-effectiveness analysis, which relates project costs to results achieved. However, this type of study requires specialized analytic/statistical skills that go beyond the scope of this manual. Nevertheless, cost issues are important to consider, since they have implications both in terms of the sustainability of project activities and outcomes and whether interventions can be successfully brought to scale. The reference list at the end of this manual includes some resources that can provide additional information on conducting cost analyses.

Below we describe the three types of evaluation in detail and present examples of each based on an illustrative project.

Process Evaluation

Process evaluation is the measurement of products and services provided by a program and the quality of those services and products. Process evaluation allows us to gain an in-depth

understanding of project implementation, including:

- number of activities carried out (often in relation to the original plan);
- quality of the activities implemented;
- reaction of the target audience (e.g., user or client satisfaction); and
- problems or obstacles encountered.

It answers the questions, “How much have we done?,” “How well have we done it?,” and “How can we improve?”

Process evaluation focuses entirely on the implementation of program activities. It does not measure how effective these activities were in producing the desired results.

The greatest benefit of this type of evaluation is its ability to identify — while the project is in full operation — the successful aspects to be continued and the deficiencies to be addressed. If program managers evaluate in a timely fashion, they can use the results to make midcourse corrections, thus increasing the chances that the program will ultimately achieve its objectives.

This type of evaluation — monitoring of program activities — is most important for the organization implementing the project and other stakeholders (e.g., local government). It is also of interest to the donor agency in that it demonstrates that the implementing agency is actively seeking to improve its services and to satisfy the needs of the intended audience.

Monitoring of Results (Outputs, Outcomes)

Process evaluation is generally easier than measuring results, especially when process evaluation involves counting number of activities completed or number of clients/participants.

Process evaluation, however, is only the first step. What we

really want to know is whether the project is making a difference. In a sense we would like to know the project's effect by measuring knowledge, attitudes, skills, behaviors and practices of the population which we are trying to help.

Successful projects have clear, realistic and measurable objectives. Monitoring of results measures the extent to which the result or desired change is achieved (or, in other words, the extent to which the objectives are met). Generally, the change in question relates to knowledge, attitudes, or practices. Monitoring of results allows us to determine if the desired change has occurred among the intended audience and, if so, how large the change is.

A second type of change involves the way in which services are provided. For example, projects may aim to increase access to services or the quality of services. In this case the evaluation may focus on the service facility as the unit of analysis (such as the number of VCT sites established in a given year).

To measure change, the evaluator must have data from before and after the intervention. Alternatively, the evaluator can establish the expected level to be achieved (in terms of absolute numbers or percentages) and then determine whether the project achieves this level in a given period of time. For example, by the end of year 1, 20% of males 15-19 will report condom use at last sex. Often however, we do not know the pre-intervention level, which is a limitation of this alternative. As a result, an increasing number of NGOs are conducting baseline (e.g., pre-intervention) assessments such as small-scale, population-based cluster surveys. They not only use findings from these assessments to set targets for key project outcomes, but also to build consensus among different stakeholders in terms of local needs and priorities.

Impact Assessment (Cause and Effect)

Impact assessment is the measurement of the health, economic status and quality of life of the target population. Impact assessment focuses on population-based measures.

Certain study designs — called experimental designs — allow us to evaluate cause-and-effect with relative precision. The most widely known of these designs is the pretest-posttest control group design, with randomization (Fisher and Foreit, 2002).

With this type of design, we are able to measure the amount of change attributable to the intervention, eliminating the possibility that confounding factors unrelated to the program influenced the results obtained. We can answer the question, “What would have happened in the absence of our program?”

In addition to experimental designs, other methodologies exist that can measure program effects. Other widely used methods include longitudinal multivariate analysis and multi-level multivariate analysis. Using appropriate statistical techniques, the evaluator can measure the extent of change that has occurred. Moreover, he/she can identify the relative importance of different factors — including exposure to the program intervention — to explain the observed change. However, due to the large samples and complex statistical analysis required, this type of approach may not be practical for NGOs working in HIV/AIDS.

Following, we present an illustrative project and explain how the three types of evaluation would apply to it.

Example of the Three Types of Evaluation Applied to an Illustrative NGO Project

The illustrative project involves commercial sex workers, or CSWs (Project 9 in Box 4). For the sake of this illustration, let's say that the objective of this project is to increase knowledge of the correct use of condoms among participating CSWs.

The intervention intended to achieve this objective is a series of workshops addressing the correct use of condoms, which will be carried out in places established especially for CSWs (such as the “safe haven” locations that exist for CSWs in a number of countries). Each participating CSW is expected to attend a workshop for one hour. The evaluator measures “correct use of condom” based on the ability to complete three actions:

1. Open the packet without using teeth or scissors (using the fingertips);
2. Remove the air from the tip of the condom; and
3. Unroll the condom using the “dildo” (anatomical model) to the base of the erect penis.¹

Example 1: Process Evaluation

With process evaluation, we want to answer two questions: “How much have we done?” and “How well have we done it?” We evaluate the first question in quantitative terms (judged against the original implementation plan). For example, if the work plan called for 10 workshops with a total of 160 CSWs, but only eight workshops (of the 10) were carried out with 128 CSWs, then the program implemented only 80% of what was anticipated.

¹ A fourth important action, verifying the expiration date, is not mentioned here due to difficulty in accurately measuring this action during a demonstration of condom use with a model.

To answer the second question, we use quantitative and qualitative techniques to assess “how well we have done it” in terms of implementing the eight workshops. In Chapter 4, we present various quantitative and qualitative techniques used in process evaluation, such as structured interviews with participants, direct observation of workshop facilitators (to evaluate the quality of the presentation), focus groups, and other techniques appropriate to process evaluation.

Example 2: Monitoring of Results (Change)

We want to determine if the project has achieved its objective, which in this case is:

- to increase the correct use of condoms among commercial sex workers.

It is important to define how to measure the indicator (known as the “operational definition”). More specifically, we have to define what we mean by the term “correct use of condoms.”

We know in advance that we will never measure correct condom use under natural conditions (e.g., during an actual sexual act between the CSW and her client). Rather, we will measure it through a demonstration with an anatomical model of an erect penis. As mentioned above, we will define the correct use of a condom by completion of three actions:

1. Open the packet without using teeth or scissors (using the fingertips);
2. Remove the air from the tip of the condom; and
3. Unroll the condom to the base of the erect penis, using an anatomical model.

If the CSW performs these three actions correctly during a demonstration using the model, the evaluator concludes that she knows how to correctly use a condom.

It is best to determine the desired levels of change to be achieved **before** the project is carried out. For example:

Desired levels

- At the end of the workshop, 90% of CSWs will be able to demonstrate correct condom use with an anatomical model.
- One month after the workshop, 80% of the CSWs that participated in the workshop will be able to demonstrate correct condom use with an anatomical model.

Box 5: Monitoring Results (Change)

Illustrative Results (Expressed as Indicators)

Percentage of CSWs that correctly perform three actions associated with correct condom use:

	Before the workshop	At the end of the workshop	One month after workshop
	%	%	%
Open packet correctly	80	90	88
Remove air from tip	40	85	68
Unroll condom onto model	60	85	65
All three actions	49	80	60

Conclusion: 80% of the CSWs demonstrate the correct use of a condom with the anatomical model at the end of the workshop. However, one month after the workshop, the percentage of the CSWs capable of demonstrating all three actions drops to 60 percent.

Monitoring results generally requires measuring the indicator(s) of interest before and after the intervention. The example (See Box 5.) shows the type of data that typically result from monitoring. Note that, in this example, the evaluation design measures the correct use of the condom at two points after the intervention: at the end of the workshop and one month after. The rationale for doing so is explained later.

This type of evaluation — monitoring of results — satisfies our desire to know if the desired change has occurred or not. By including a follow-up assessment some duration (in this case, one month) after the workshop, it is also possible to assess whether the desired results are sustained. For many projects, the approach illustrated in Box 5 is adequate, especially if the project does not have sufficient staff or budget to carry out a more ambitious evaluation work.

However, it is possible that the changes observed are due to other factors (not necessarily or not exclusively due to our project). For example, the CSWs might have listened to messages on condom use through the mass media or from peer educators from a different agency working in the same area. To eliminate the possibility that the observed changes result from factors other than our own program, we would need to conduct an impact assessment using a design that would demonstrate cause and effect.

Example 3: Impact Assessment (To Demonstrate Cause and Effect)

The classic example for demonstrating cause and effect is the experimental design. Although several variations exist, the best known is the pretest-posttest randomized control group design (Fisher and Foreit, 2002). The evaluator must randomly assign participants in the study to one of two groups: the experimental (or treatment) group (which receives the intervention) or the control group (which does not). In choosing a control

(comparison) group, it is ideal to select individuals who are generally similar to individuals in the experimental group, with one major exception: “exposure” to the intervention. In our example of the project with CSWs, the evaluator could randomly assign participants to the experimental and control groups.²

Box 6: Example — Impact Assessment

Illustrative Result (Expressed as an Indicator):

Percentage of CSWs who correctly perform all three actions

	Before the workshop	One month after the workshop
Experimental Group	33%	65%
Control Group	35%	38%

Conclusion: The workshop was responsible for an increase in correct condom use, given a significant increase in the experimental group that did not occur in the control group. The small change in the control group indicates what happened for other reasons — in the absence of the intervention.

Let’s assume in our illustrative project with CSWs that we are concerned with the correct use of the condom, not so much at the end of the workshop but a month after the workshop, to better assess the true effect of the workshop on the CSWs’

² Random assignment implies use of special techniques that remove bias from the formation of groups, such as assigning every second woman to the control group or using a table of random numbers. For ethical reasons, evaluators often arrange for members of the control group to have the opportunity to receive the same intervention (e.g., workshop) after the evaluation is completed.

knowledge. Thus, to simplify the example, we will focus only on correct use one month after the workshop and only on correct knowledge of all three actions (not each separately). *Box 6* compares the two groups (experimental and control) on the key indicator: correct use of condoms.

Experimental designs are generally considered to be the most rigorous means of evaluating impact. However, they tend to be expensive and time-consuming; and they require statistical expertise for appropriate data analysis. Even more important, they are impractical for programs that are “full coverage” (designed to reach the entire population), such as those with a mass media component.

Types of Evaluations Conducted to Date on HIV/AIDS Prevention Projects

To date, the vast majority of evaluations conducted by NGOs on HIV/AIDS projects in developing countries has consisted of process evaluation or monitoring of results. Very few organizations have conducted impact assessment without a substantial influx of resources from external sources. Of note, the National Institutes of Health in the United States is funding an ambitious program of clinical trials for behavior change on HIV/AIDS; yet these research projects differ substantially from the type of evaluation that most governments or NGOs in developing countries are prepared to conduct.

Given the complexities of evaluating behavior change in relation to HIV/AIDS, most organizations (and where relevant, their donors) are satisfied to stop at process evaluation and monitoring of results. While we encourage organizations to take advantage of opportunities that may arise to conduct impact assessments using experimental or quasi-experimental designs, most NGOs leave this type of evaluation to those with the funding, training, and statistical experience to carry it out.

Chapter 3

Designing an Evaluation

This chapter focuses on evaluation designs appropriate for NGOs working on HIV/AIDS. We will cover process evaluation and monitoring of results but not impact assessment. (Those interested in this topic should consult Fisher and Foreit, 2002, chapter 7.) Moreover, we will focus on evaluations at the program level (involving clients or program participants) and not at the population level (general population or subgroup in a country, region or city).

Steps in Designing an Evaluation

In designing a process evaluation or monitoring of results, we follow five basic steps:

1. Determine (or clarify) the project's objectives;
2. Determine the type of evaluation to be carried out;
3. Identify indicators based on the type of evaluation;
4. Determine the appropriate data source for each indicator;
and
5. Prepare an evaluation plan.

The program manager and evaluator, in collaboration with other stakeholders, should agree on these five steps **before the project is implemented**. Otherwise, the time and money invested in evaluation may not yield optimal results. It is very important to be transparent and participatory in evaluation.

Doing so increases the likelihood that the relevant stakeholders will accept evaluation results when they become available.

Step 1: Determine (or Clarify) the Project's Objectives

When we design a project, it is for a purpose, that is, with the objective of bringing about a specific behavior change. Clearly defining the objectives of a program or project may be a determining factor in obtaining (or not obtaining) financial support from a donor agency.

Although HIV/AIDS prevention programs and projects take many different forms worldwide, many have a common set of objectives, as summarized in Box 7.

Ideally, the program objectives are SMART:

- **S**pecific (State who will be doing what, when, how much and who will benefit.)
- **M**easurable (Intended accomplishments must be measurable.)
- **A**ttainable (Objectives should be realistic given the context and available resources.)
- **R**elevant (The objective should relate to the problems discovered in the needs assessment.)
- **T**ime-bound (It is important to mention the time-frame in which the objective will be achieved.)

Box 7: Common Objectives in HIV/AIDS Prevention Programs/Projects³

An increase or improvement during a given period in:

Knowledge (of):

- modes of HIV transmission
- methods of preventing transmission
- sources of condom distribution and of VCT
- correct use of condoms

Attitudes/Perceptions:

- acceptance of people living with HIV/AIDS
- attitudes towards the rights of people living with HIV/AIDS
- perception of personal HIV risk
- self-efficacy in adopting certain risk-reduction behaviors

Behavior (or practices):

- delayed age at first sex
- remaining faithful to one partner
- use of a condom in last sexual relation with a casual partner (that is, a non-marital, non-cohabiting partner)
- consistent use of condoms in sexual relations (or in sexual relations with casual partners)
- number of sexual partners (decreased)

³The objectives for this box are not “SMART.” They were not worded in a SMART manner because the exact details would vary from project to project. An example of a SMART initiative would be: to increase the percentage of the general public with anonymous means of receiving information and assistance from 5 percent to 20 percent by 2006.

Some NGOs have difficulty formulating program objectives. Box 8 presents one common pitfall.

Box 8: A Common Error in Formulating Objectives

One common error in formulating program objectives is to describe the activity to be carried out rather than the objective (result) to be achieved.

Correct (specifies desired change):

- 90% of commercial sex workers participating in the workshop will know how to correctly use a condom one month after they have been trained.

Incorrect (focuses on activity; does not indicate desired change):

- Conduct 20 educational talks on condom use with CSWs participating in the program.

Step 2: Determine the Type of Evaluation to Be Carried Out

As mentioned earlier, the large majority of evaluations of HIV/AIDS projects consist of:

- process evaluations: to measure how much we have done and how well we have done it; or
- monitoring of results (change): to determine the extent to which the objectives have been achieved.

Frequently the evaluation of a given project will include both components. Whereas an evaluation can include one without the other, we strongly recommend conducting both types.

In sum, the second step is to decide whether to carry out a process evaluation, monitoring of results, or both. (Should

circumstances permit, the project manager and evaluator might also opt for an impact assessment.)

Step 3: Identify Indicators for Each Type of Evaluation

In this step, we identify the indicators needed to evaluate implementation of the project (process evaluation) and/or the results obtained (monitoring of results).

To determine the most appropriate indicators for a given evaluation, we ask, “What do we want to learn from this evaluation?”

Box 9: What Is an Indicator?

An indicator is a measure that describes a behavior, concept, or phenomenon. It does not have to capture all of the aspects of the phenomenon to be measured, but it should give an indication of this. For example, if we want to measure "acceptance of people living with HIV/AIDS," the concept is complex. We attempt to get some measure (or indicator) that captures this concept. For example, many surveys measure the percentage of people who:

- would be willing to share a meal with a person they knew had HIV or AIDS
- would be willing to care for a relative who became sick with the virus in their own household
- believe that a female teacher that has the AIDS virus should be allowed to continue teaching in the school

Each of the above is an indicator of acceptance of people living with HIV/AIDS. The concept of acceptance is much broader than these three points; however, these indicators allow us to attempt to quantify this phenomenon. An indicator is expressed as an absolute number, a percentage, a rate (prevalence or incidence), or “yes/no.”

Process evaluation can use both quantitative and qualitative data. Quantitative indicators measure numbers, percentages, averages, and other statistics. We may also want to track certain aspects of the program using qualitative data. Both types of data can yield valuable insights into the workings of a program. For example, an evaluator could assess the level of satisfaction with a workshop by administering a structured questionnaire to measure the percentage of participants who rated the presentations to be understandable, interesting, or informative (quantitative). Alternatively, the evaluator could convene the participants in a focus group to learn about their attitudes toward different aspects of the workshop (qualitative data). Both methods have advantages and disadvantages, but both provide potentially useful feedback on the dynamics of the workshop.

Monitoring of results requires quantitative data to measure the amount of change achieved. In Box 7 we presented a list of common objectives for HIV/AIDS prevention projects. In Box 10, we convert these objectives into indicators that can be measured in an evaluation.

In Chapter 5, we outline the indicators corresponding to the most common project interventions. A summary of the list for all interventions appears in Appendix A.

It is important to note that because of their empowerment focus, NGOs recognize that social norms, myths, and perceptions strongly influence the end behaviors they are trying to promote (e.g., using condoms, limiting the number of sexual partners, or practicing abstinence). In addition to the indicators listed in Box 10, a number of NGOs measure indicators that reflect additional precursors to behavior change. Examples of such indicators are the percentage of the intended audience who believe they are at risk of getting HIV, the percentage that feel confident they could avoid having sex with a partner if they wanted to, and the percentage who feel

confident that they could convince their sexual partner(s) to use condoms.

In selecting indicators, keep in mind that:

- There must be at least one indicator for each result or objective to be measured.
- If the indicator measures a phenomenon that does not change in the short term (for example, in less than a year), it is not worthwhile to measure the indicator every 12 months; the evaluation should measure it at longer intervals. (It might be possible, however, to identify some process issues that can be measured more frequently, and that could be used to gauge progress towards the intended result.)
- If the indicator is not useful for programmatic purposes (that is, to improve the program or assess change), the evaluator should drop it from the evaluation.

Box 10: Common Indicators for Monitoring Results in HIV/AIDS Prevention Programs/Projects

Knowledge:

Percentage of the intended audience that knows:

- three modes of HIV transmission
- three methods of preventing HIV transmission
- a source of condom distribution (alternatively, the number of sources known)
- three actions for correct condom use
- three ways HIV can be transmitted from mother to child

Box 10 (cont.)

Attitudes/Perceptions:

Percentage of the intended audience that:

- would be willing to share a meal with a person they knew had HIV or AIDS
- would be willing to care for a relative who became sick with the virus in their own household
- believe that a female teacher that has the AIDS virus should be allowed to continue teaching in the school
- perceive themselves to be at risk for HIV/AIDS
- feel confident they could avoid having sex with a partner if they wanted to
- feel confident they could convince their sexual partner(s) to use condoms

Behavior (or practices):

Percentage of the intended audience that:

- practice abstinence
- have had only one partner in the last 12 months
- used a condom at last sex with a non-regular partner (if sex workers, with their client or partner)
- use a condom for each sexual relation (“always”)

Step 4: Determine the Appropriate Data Source for Each Indicator

The evaluator must have a source of data for every indicator. In some cases, the data are already available (e.g., routine service statistics); it is only a question of analyzing them. In other cases, the evaluator must design and conduct a study to obtain the data.

The majority of evaluations utilize quantitative data to answer the question, “How much?” The most common quantitative data sources are:

- structured interviews (surveys); and
- service statistics (and other program data).

Quantitative data are not, however, as rich in detail on attitudes, opinions, and values as qualitative research techniques, such as:

- focus groups;
- in-depth interviews;
- direct observation;
- ethnographic observation; and
- other qualitative techniques.

Chapter 4 explains the different sources of data in more detail. Chapter 5 on “HIV/AIDS Illustrative Projects” gives examples of indicators and their data sources for illustrative HIV/AIDS projects (listed previously in Box 4).

Step 5: Prepare an Evaluation Plan

An evaluation plan summarizes the decisions made regarding the four steps described above, including:

- The objectives of the project (for the specific target population)
- The type of evaluation to be carried out
- The indicators to be measured
- The appropriate data source for each indicator.

Box 11 presents the various topics that should be included in the evaluation plan.

Box 11: Contents of a Illustrative Evaluation Plan

Brief Project Description:

- Background
- Implementing organization
- Intended audience
- Planned interventions

Objectives of the Project:

Objectives of the Evaluation:

Evaluation Methodology:

- Type of evaluation
- Indicators to measure each objective
- Data collection method(s) (e.g., surveys, interviews, focus groups, service statistics)
- Sample size
- Methods of sampling (or selecting participants)
- Data collection, processing, and analysis (how, where, who)

Resources:

- Material (supplies and equipment)
- Human
- Financial
- Transportation and logistics

Anticipated Use of Results:

- To improve the project mid-course
- To plan future projects
- To guide decision-making

Timetable:

Budget:

Some Suggestions for Evaluating Projects with Multiple Components

Programs or projects that are complex — with multiple audiences, objectives, and programmatic approaches — present special challenges to the evaluator. In this final section of Chapter 3, we offer suggestions for dealing with these challenges. Although these suggestions relate primarily to the presentation of evaluation findings, it is useful to keep them in mind during the design of the evaluation.

Programs often have:

- more than one intended audience;
- more than one objective to achieve; and
- more than one programmatic approach (intervention strategy).

What is the best way to handle this situation in presenting the results? We suggest that once the evaluation is completed and data are available, the evaluator present the results in a logical format that the reader can easily follow. Three options include the following:

Option 1: Presentation by Subgroup (Different Intended Audiences)

Example: Suppose we are evaluating an educational activity on the correct use of condoms among two specific subgroups: commercial sex workers and men who have sex with men.

Suggestion: Organize the evaluation and presentation of the results in parallel form for the two subgroups (especially if the objectives to be achieved are different in each population). Specifically, present the findings for the first group in their entirety (without mention of the second group). Then do the

same for the results pertaining to the second group. Finally, discuss the similarities and differences in the results for the two groups and make recommendations specific to each.

Option 2: Presentation by Methodology Used

Example: Assume we are conducting an evaluation of an educational activity on the correct and consistent use of condoms among commercial sex workers and men who have sex with men (same as above), using both structured interviews and direct observation.

Suggestion: Organize the evaluation (and subsequent presentation of results) according to methodology used to collect the data, for example: structured interviews and direct observation.

Indicate what data are obtained from each data source and how they contributed to the evaluation of the objective(s) for each of the target populations, comparing and contrasting the findings from the different methodologies used.

For example, in this case it might work well first to conduct structured interviews that measure knowledge obtained from the educational activity and then to test participants on their ability to correctly use a condom, based on observation using an anatomical model. The evaluator should explain the type of information to be obtained from the first data source (structured interviews) and from the second source (direct observation), indicating how they both contribute to meeting the objectives of the evaluation.

Option 3: Presentation by Type of Intervention

Example: Suppose we are evaluating a social marketing program for condoms that utilizes two main channels to reach the intended audience: (1) radio and (2) distribution of condoms and printed material in bars. The evaluation uses surveys and focus group data.

Suggestion: Organize the evaluation according to the different interventions carried out. Indicate what type of information is obtained from the intervention carried by radio; then describe the information that will be obtained from the intervention in bars, specifying how each contributes to the objectives of the program.

Chapter 4

Methodologies for Collecting Data

Two categories of methods exist to collect data for project evaluation: quantitative and qualitative. The same methods utilized in social research are equally applicable in program evaluation.

What Is the Difference between Quantitative and Qualitative Methodologies?

Quantitative techniques describe different aspects of the project in terms of numbers (e.g., absolute numbers, percentages, means), such as the number of pamphlets published or the percentage of the target population that knows three methods of HIV transmission. In the case of quantitative methodologies, the evaluator defines what he/she considers important to measure and focuses on obtaining valid data on these points.

The most common quantitative techniques used in program evaluation are:

- Structured interviews (with clients, participants, service providers, and others) and
- Service statistics (or other program data).

Qualitative techniques differ in two important aspects. First, they generally do not try to describe the project with numbers, even though at times, evaluators may quantify phenomena

observed using qualitative techniques. Second, qualitative techniques allow respondents to discuss the project in their own terms of reference and to identify issues that are important for them, which may differ from the evaluator's perspective. Although a discussion or observation guide exists, the people who participate in the evaluation have much more freedom to express their ideas or describe the situation according to their own point of view.

The most common qualitative techniques used in program evaluation are:

- focus groups;
- in-depth interviews; and
- observation (direct observation, mystery client, or ethnologic techniques).

Box 12 summarizes key differences between quantitative and qualitative techniques. The purpose of this chapter is to describe quantitative and qualitative techniques in further detail. For each technique, we have included a table that describes the main steps in using it.

Quantitative Evaluation Methods

Structured Interviews

The typical survey uses structured interviews. A trained interviewer poses a series of preestablished, carefully worded questions to the respondent. He asks the same questions to every respondent, using the same wording in the same order.

In structured interviews, the questions are often closed-ended with pre-coded responses. The evaluator can anticipate most of the responses respondents will give and lists them on the

questionnaire. The instrument may also contain some open-ended questions, which are coded after all the responses are in.

Box 12: Some Differences between Quantitative and Qualitative Methods⁴

QUANTITATIVE METHODS	QUALITATIVE METHODS
Describes "how many?" or "how much?"	Describes "how?" and "why?"
Uses predominately closed-ended questions	Uses predominately open-ended questions ⁵
Provides numerical data and statistics that facilitate similar interpretation by evaluators	Provides data on perceptions, beliefs, and values, which can be interpreted differently by different evaluators
Requires large samples, preferably selected at random	Permits more limited samples, generally not selected at random
Requires staff with experience in statistical methods	Requires expertise in qualitative data analysis
Results can be generalized to the target population	Results cannot be generalized and are only indicative of a segment of the population
Yields more superficial responses to sensitive topics, (e.g., sexual behavior)	Offers more in-depth responses on sensitive topics, (e.g., sexual behavior)

⁴ Adapted from the following source: Debus, Mary. *The Handbook for Excellence in Focus Group Research*. Washington, DC: Academy for Educational Development (AED).

⁵ With the exception of observation.

For example:

Closed-ended question:

Do you think you are at risk of getting HIV/AIDS?

1. Yes

2. No

Open-ended question:

Why do you think you are at risk of getting HIV/AIDS?

The self-administered interview is a variation on the structured interviews, in which the respondents complete the interview themselves. The instrument consists of a series of questions, which can be closed (pre-coded) or open-ended.

Self-administered questionnaires only work among populations that know how to read and write. Even so, make the wording simple and the instructions clear. A problem with self-administered questionnaires is that the respondents may leave questions blank or fail to follow the instructions correctly (e.g., in terms of skip patterns). Furthermore, in many countries, the culture for mail-back questionnaires does not exist.

Box 13: Steps in Carrying Out an Interview Using a Structured Questionnaire

Plan the interview:

- Review the study, including the survey instrument by an IRB if necessary.
- Define the protocol of the evaluation, including:
 - the objective of the interview
 - the target population
 - calculation of the sample size
 - timeline
- Develop the questionnaire.
- Review the questionnaire with experts in survey research, as well as program stakeholders, and incorporate revisions.
- Select and train the interviewers.

Carry out the field work:

- Conduct a pre-test in the field (among respondents similar to the population to be interviewed).
- Modify the instruments based on the pre-test.
- Coordinate logistical aspects for the field work.
- Collect the data.

Revise, process, and analyze the data:

- Review the questionnaires while the interviewers are still on location.
- Code the data.
- Enter the data into the computer (with a program such as Epi-Info or SPSS).
- Design an analysis plan.
- Process the data (with a program such as Epi-Info or SPSS, or by hand).
- Prepare the tables of results, according to the plan of analysis.

Produce the report and disseminate results:

- Prepare a final report.
- Share the results with those responsible for the project and other interested parties.

Service Statistics

The term “service statistics” is used in a broad sense to include any type of information that can be obtained from the project records, including:

- logs of client visits, registration books;
- program records on activities carried out;
- cardex files; and
- inventories of materials produced and distributed.

The major advantage of services statistics is that the information is routinely collected as part of the operation and does not represent an additional cost, especially if the information system is designed to produce data on the indicators of interest to the evaluation.

In terms of disadvantages, (1) it is only possible to obtain information on variables existing in the system, and (2) the data are not always reliable or complete.

***Box 14: How To Use Service Statistics
for Evaluation Purposes***

- **Define the indicators to be measured from service statistics**
- **Determine if the indicators are measurable by data existing in the system**

If the data exist in the information system:

- Ensure the quality (accuracy, reliability) of this data and, if necessary, find ways to improve the quality, such as more training and supervision
- Process the data (indicators) of interest for the evaluation manually or by computer
- Present the data in a form that is easy to understand (e.g., graphs)

If the data do not exist in the information system:

- Redesign the form to yield the data of interest
- Obtain authorization to incorporate the changes into the information system
- Train appropriate staff in the use of the revised/new instrument
- Ensure the quality (reliability) of the data collection
- Process the information of interest to the evaluation manually or by computer
- Present the information in a form that is easy to understand (e.g., graphs)

Qualitative Methods

Qualitative methodologies provide depth to a given evaluation by exploring the “why” in a given situation, attempting to understand the context for certain behaviors and the emotions of the persons involved.

For example:

A quantitative evaluation tells us what percentage of commercial sex workers has dropped out of a program designed for them. A qualitative evaluation explores the reasons **why** the commercial sex workers have dropped out of the program.

The qualitative techniques most frequently used for program evaluation include the following.

Observation

This technique consists of observing a specific event (such as an educational talk about HIV/AIDS given to factory workers) or an activity that is frequently repeated (such as a counselor responding to calls on an information hot line). Three main forms of observation are:

- direct observation (with a guide or checklist)
- systematic observation by a mystery (or simulated) client
- ethnographic observation

Direct Observation (with a Guide or Checklist)

This technique is frequently used to evaluate activities of key importance to the project, for example the technical competence of clinical staff or educators/facilitators. It is possible to reduce the subjectivity by creating a checklist or

guide that outlines a series of actions that the person under observation should be doing and the criteria that constitute a “correct action.” Box 15 provides examples of checklist items for direct observation.

***Box 15: Examples of Checklist Items
for Direct Observation***

Suppose the evaluation focused on medical and paramedical staff to determine if they observe universal precautions for HIV prevention at their work site. The checklist could include the following aspects:

- wash their hands before examining the client;
- wash their hands after examining the client;
- use gloves when handling secretions and bodily fluids;
- dispose of needles adequately and avoid accidental pricks; and
- discard all used supplies or place them in the sterilizing apparatus.

In its simplest form, an observer attends an activity, “assesses” if the person carrying out the activity does it adequately, and offers suggestions for improvement. However, this type of evaluation can be very subjective if criteria for evaluation are lacking or are not easy to observe.

This research technique requires highly qualified personnel and sufficient time to observe the different staff working on a given project. The costs can be high if the action to be observed happens infrequently or if the sample is large. The following table summarizes important considerations when carrying out direct observation.

Box 16: How to Carry Out Direct Observation

Plan the evaluation:

Define the evaluation protocol, including:

- Objective of the observation
- Intended audience
- Sample size
- Timeline

Define the specific procedures to be observed during the performance of the work (e.g., greeting upon entry into the clinic, privacy during counseling, waiting time)

Design the instrument (checklist) that includes each specific action considered important to successful completion of the task

Select and train one or more persons (in the above example, those with clinical training) capable of evaluating the professional performance of the staff observed

Carry out the field work:

Arrange all logistical aspects of the evaluation

Review the instrument with other persons with expertise in this area

Conduct a pre-test of the instrument with people similar to those who will be evaluated

Carry out the direct observation according to a protocol

Process and analyze the data (manually or by computer)

Create tables and graphs that demonstrate results in a readily understandable format

Prepare a final report of the results

Present and discuss the results with those responsible for the project and other interested parties

Systematic Observation by a Mystery (or Simulated) Client

This technique draws on the approach used in the commercial marketing sector, known as the “mystery shopper.” A member of the evaluation team disguises him/herself as a customer or client and observes the level of service that the project staff provide when they are unaware that they are under observation. Box 17 summarizes how to carry out an assessment using mystery clients.

Box 17: How to Carry Out an Observation Using Mystery Clients

Plan the evaluation:

Define the evaluation protocol including:

- objective
- intended audience
- sample (how many people and how they will be selected)
- time line for collecting the information

Establish a “profile” of the mystery client and a storyline that can be used consistently at every visit. (For example, a gay man in his late 20s, with a high school education, has not used condoms and now believes that he could be infected with HIV because of his cough and loss of weight.)

Select and train a mystery client who fits the established profile.

Identify the items or actions to be observed (e.g., greeting at entry into clinic, privacy during counseling, waiting time under 15 minutes).

Design an observation instrument that includes each action considered important.

Box 17 (cont.)

Carry out field work:

- Pretest the instrument in the field and modify it as necessary.
- Carry out the observation according to the work plan.
- After each observation (that is, visit to obtain service or product) register the information using one of the two alternatives:
 - a. The mystery client leaves the site and completes the form based on what he/she remembers.
 - b. The mystery client leaves the site and another member of the team interviews the mystery client on each key item or action; interviewer fills out the form.

Process and analyze the results:

- Process the data manually or by computer.
- Identify the tables and graphs that show the results in a readily understandable form.

Prepare reports and disseminate results:

- Prepare the final report of results.
- Present and discuss the results with those responsible for the project and other interested parties (e.g., health staff that participated in the evaluation).
- Develop action plans for addressing deficiencies identified during the assessment.

The main advantage of this technique is to be able to observe what happens under “normal circumstances,” when the personnel are not conscious of being observed.

One approach is to have the observer commit to memory the list of items to observe; then directly after the observation

(and out of sight of other clients), he/she completes a form containing this series of items. Alternatively, upon completion of the observation, an “interviewer” can question the observer about the series of points and record the information on a similar form.

This methodology has various limitations and is used less frequently than direct observation for evaluation.

1. The observer may not accurately remember what he/she observed while visiting the project or service.
2. If one of the service providers recognizes the mystery client, it is necessary to terminate the observation and eliminate the case from the analysis.
3. Many program administrators do not feel comfortable with a methodology that seems like “spying.” One way of addressing this criticism is to inform the health system staff that during the upcoming months, a mystery client study will be carried out in some facilities, without specifying the sites or the dates of the visit.

Ethnographic Observation

This technique consists of observing the sociocultural context and human relationships that influence a given behavior or event in a given population over a period of time. It is useful in describing and explaining the behaviors and forms of interaction occurring in a given population relative to a given phenomenon or event.

For example:

Ethnographic observation of truck stops along a major highway could be useful in understanding the dynamics of behavior between truck drivers and commercial sex workers along the route.

In the “classic” ethnographic observation, the ethnographer lives and works for an extended period of time with a group or community under study, observing their interaction through participation in the community life of this group. Given the large amount of time required for this methodology, it is not practical for evaluation purposes. Thus, we do not describe the steps to carry out this type of observation in this manual.

In-depth Interviews

These interviews are carried out in person, using a discussion guide with a logical sequence of questions; however, the interviewer has considerable flexibility in conducting the interview. He/she is free to discuss topics from the guide as they spontaneously arise, as well as to explore other topics not mentioned on the guide but relevant to the subject. Box 18 presents key issues related to conducting in-depth interviews.

Box 18: How to Carry Out an In-depth Interview

Plan the evaluation:

- Define the evaluation protocol, including:
 - objective
 - intended audience
 - sample (how many people and how selected)
 - time line
- Be sure to obtain IRB approval if needed.
- Develop a discussion guide (topics to be covered).
- Review the guide with input from experts in this area.
- Select and train one or more interviewers regarding:
 - objectives of the evaluation
 - in-depth interview techniques
 - basic facts about HIV/AIDS

Box 18 (cont.)**Carry out the field work:**

- Make logistical arrangements.
- Carry out a pretest in the field with people similar to those who will be interviewed.
- Carry out a predetermined number of interviews, taping each of the sessions.

Transcribe and analyze the data:

- Transcribe the recorded information verbatim (translating from the local language to the evaluator's language if necessary).
- Design a plan for analyzing the information (i.e., a list of topics to cover).
- Apply qualitative data analysis techniques to identify the principal conclusions from the in-depth interviews.

Prepare report and disseminate results:

- Prepare a final report.
- Share results with those responsible for the project and other interested parties.

In-depth interviews reveal a great deal about the attitudes, motivations, and mind-set of the respondents. Moreover, they allow for topics of interest to the respondents to surface, which the evaluator had not thought to include.

The value of the results from in-depth interviews depends in large part on the skill of the interviewer in posing questions and probing for responses without inadvertently biasing the respondents' answers.

It is advisable to tape the interviews and to transcribe the results to have a permanent record of the discussions. However,

this process requires substantial investment of time and money, which is one of the disadvantages of this methodology. Moreover, it is essential to have well-trained personnel conduct the data analysis in a systematic way to reduce the bias of subjectivity. Another disadvantage of taping the session is that it may make some respondents ill at ease, thus influencing the quality of their responses.

Focus Groups

This qualitative method brings together a group of people from the community or target population to discuss a topic of interest to the evaluator.

Focus groups provide in-depth insights into the attitudes, beliefs, values, and perceptions of a given group on a specific subject (e.g., the “focus” of the discussion), although the results are not necessarily representative of the larger population. When done well, focus groups provide an incredible richness of information on what is in the hearts and minds of specific groups (e.g., the intended audience).

Prior to convening the actual discussion groups, the evaluator creates a discussion guide of the topics to be covered, often starting with a general question that creates an atmosphere of trust, and then moving to more specific, often sensitive issues. In a good focus group, the conversation appears to flow naturally and the moderator is flexible in allowing participants to discuss what is on their mind. Yet unbeknownst to the participants, the conversation follows the topics on the discussion guide. The moderator does not have to respect the order of the questions on the guide, but should try to cover all or most of them during the session.

It is important to have a moderator and an assistant (observer) trained in focus group methods and with demonstrated skill in conducting focus groups. With respect to evaluation, focus

groups can serve several useful purposes. They can guide the design of quantitative studies (for example, baseline and follow-up surveys) in terms of topics to cover and appropriate language (e.g., local terms) to use for specific concepts. They yield feedback on the functioning of a given program from its beneficiaries. They can also provide insights into the reasons a program does (or does not) have the desired impact. Box 19 presents important issues when carrying out focus groups.

Box 19: How to Carry Out Focus Groups

Plan the evaluation:

- Define the evaluation protocol, including:
 - objective
 - intended audience
 - sample (how many people and with what characteristics)
 - time line
- Develop a discussion guide (topics to be covered)
- Review the guide with input from experts in this area
- Select and train one or more moderators with regard to:
 - objectives of the evaluation
 - indepth interview techniques
 - basic facts about HIV/AIDS

Carry out the field work:

- Carry out at least one practice focus group among people with characteristics similar to the population to be evaluated before actually collecting data.
- Conduct the pre-determined number of groups, tape recording the sessions and taking notes.

Box 19 (cont.)**Transcribe and analyze the data:**

- Transcribe the recorded information verbatim (translating from the local language to the evaluator's language if necessary).
- Design a plan for analyzing the data (i.e., a list of topics to cover)
- Apply qualitative data analysis techniques to identify the key themes from the focus groups.

Prepare reports and disseminate results:

- Prepare a final report.
- Share results with those responsible for the program and other interested parties.

Important aspects to using qualitative research methodologies include asking why, knowing how to listen, and probing for answers (See Box 20.).

Box 20: Three Keys to Successful Qualitative Research⁶

The Art of Asking “Why?”

The experienced qualitative evaluator will be careful to:

- ask in a neutral manner;
- avoid leading the respondent;
- ask only one question at a time; and
- note verbal and nonverbal clues of confusion or evasiveness from the respondent.

The Art of Listening:

- Active listening is closely related to ***empathy***, one’s ability to identify with another in terms of the way that person would feel or act.
- The ***way*** things are said may reveal more of the intended meaning than the words that are spoken.
- Good listening requires hearing what is meant as well as what is said. This means picking up on nonverbal clues — indicators of anxiety and uncertainty, of confidence and assertiveness. Hesitations, silences, and variations in word choice are also relevant.

Use a Creative Process of Investigation:

A high level of creative thinking must be applied to each new situation if the qualitative research process is to be truly successful.

⁶ Source: Debus, M. et al. *The Handbook for Excellence in Focus Group Research*. Washington, DC: Academy for Educational Development (AED).

Example of the Use of Qualitative and Quantitative Methods: Evaluating Client Satisfaction

Many programs are interested in knowing if the services provided are fulfilling the needs and expectations of their clients (participants), and how they could improve them to have a greater impact. Below we present two methods of achieving this aim, one quantitative and one qualitative. The following boxes give examples of a structured questionnaire and focus group guide.

Box 21: Structured Questionnaire for Clients of the VCT Service

1. How did you learn about this VCT service? Mark relevant answer(s):

1. ___ referred by a doctor or other health service
2. ___ newspaper advertisement
3. ___ radio advertisement
4. ___ from a friend or relative
5. ___ Other, explain: _____

2. How many minutes did you have to wait from the time that you arrived at the clinic until someone attended to you?

___ minutes (if it is more than an hour, write the total time you waited in minutes)

3. Did the amount of time you waited seem:

1. ___ too long?
2. ___ acceptable?

Box 21 (cont.)

4. How did the staff treat you during your visit today? Mark the answer that best describes your opinion with respect to the following staff members.

a. Receptionist:

1. Excellent 2. Good 3. Fair 4. Poor

b. Counselor/service provider (doctor, nurse, counselor, social worker, etc.):

1. Excellent 2. Good 3. Fair 4. Poor

c. Laboratory technician:

1. Excellent 2. Good 3. Fair 4. Poor

5. In the pre-test counseling session, did they tell you the following:

a. That the test would require drawing blood?

1. Yes 2. No

b. How to assess your personal risk?

1. Yes 2. No

c. When you would get your results?

1. Yes 2. No

d. What the results mean?

1. Yes 2. No

e. How to protect yourself from HIV transmission?

1. Yes 2. No

f. How to correctly use a condom, using an anatomical model?

1. Yes 2. No

Box 21 (cont.)

- 6. With regard to the person who counseled you:**
- a. Did he/she seem:**
1. friendly 2. unfriendly?
 - b. Did he/she treat you with respect?**
1. Yes 2. No
 - c. Did He/she give you the opportunity to ask questions?**
1. Yes 2. No
 - d. Did he/she show interest in you as a person?**
1. Yes 2. No
- 7. In regard to the counseling:**
- a. The information provided was:**
1. too much 2. just right 3. too little
 - b. The explanations were:**
1. clear 2. unclear
- 8. On what topic(s) would you have liked to have more information?**
- 9. Were you left with any doubts or questions at the end of the counseling session?**
 Yes No
- 10. What doubts did you have?**

Box 21 (cont.)

- 11. Did you learn something new in the counseling session?**
1. ___ Yes 2. ___ No
- 12. Did one of the clinic staff ask for your permission to take a blood sample for the HIV test? (Alternatively: Did you sign a consent form before the staff took your blood?)**
1. ___ Yes 2. ___ No
- 13. How long did you wait for your test results?**
_____ hours (or _____ days)
- 14. Did the staff tell you that all information about your visit and your test results would remain confidential (for personal use only), and the staff would not have the right to share this information with other people?**
1. ___ Yes, they told me. 2. ___ No, they didn't tell me.
- 15. How certain are you that the information about your visit will remain confidential?**
1. ___ very sure 2. ___ somewhat certain
3. ___ not very certain

Why?

Box 21 (cont.)

16. With respect to the post-test counseling, did they inform you about the following:

a. The meaning of the results?

1. ___Yes 2. ___No

b. The difference between a person who is HIV positive and a person with AIDS?

1. ___Yes 2. ___No

c. Places where one can receive treatment?

1. ___Yes 2. ___No

d. Alternatives one has for treatment?

1. ___Yes 2. ___No

e. Ways of protecting one from HIV infection?

1. ___Yes 2. ___No

f. Demonstration of the correct use of condoms on an anatomical model?

1. ___Yes 2. ___No

17. Were you given any information (e.g., brochures) about HIV/AIDS or the HIV test to take with you?

1. ___Yes 2. ___No

Box 21 (cont.)

18. After you got your results:

a. Was the information provided:

1. ___ too much 2. ___ just right 3. ___ not enough

b. Were the explanations:

1. ___ clear 2. ___ unclear

19. What topics would you have liked to have more information on in the post-test counseling session after you got your results?

20. Overall, what did you think about the counseling and testing service that this clinic provides?

1. ___ excellent 2. ___ good 3. ___ fair 4. ___ poor

21. What suggestions would you have to improve the counseling and testing service at this clinic?

Thank you for your time and answers!

Box 22: Focus Group Guide

1. (Ice breaker) Tell me a little about yourself: your favorite sport or pastime.
2. How did you learn about this VCT service?
3. Before coming to this clinic/center/hospital today, what had you heard about this facility?
4. Did you have any fears or hesitations about coming today to get tested? Why?
5. What comments do you have about the procedures used to draw blood?
6. What did you think of the counseling session before they took your blood? Why? And what about the session after they gave you your results? (Probe: What was good about the session? What could be improved?)
7. What do you think about this facility in terms of its appearance? Equipment?
8. How did the staff treat you? (Probe for details)
9. Would you recommend this clinic to other people who want to get tested? Why?
10. What would you do to improve the VCT services provided at this facility?
11. Were you given any information brochures about HIV/AIDS or the HIV test to take away with you? If so, were they useful?

Box 22 presents an example of a focus group discussion guide.

Methodological Note about Focus Groups

The people who use VCT services may want to remain anonymous and avoid interacting with others receiving the same

services. In this case, it is advisable to use a structured questionnaire (as illustrated in Box 21) or a series of in-depth interviews (administered individually). However, the occasion may arise when the evaluator can conduct a focus group among clients of the service to get their feedback on VCT services.

Linking the Type of Evaluation to the Appropriate Data Source

In this manual we have treated each topic separately: type of evaluation, indicators for process evaluation and monitoring, quantitative and qualitative data sources. The challenge at this point is to decide — based on the type of evaluation to be conducted — what indicators will be most useful in evaluating a specific project and what data sources will be necessary to obtain the information.

In terms of monitoring results, the choice of indicators will depend on the objectives of the project. With regard to process evaluation, the choice will depend on what the program planners judge to be most important aspects of implementation that they should track.

There is no “recipe” that can be followed step by step; rather, the evaluator follows some general guidelines and adapts them to the specific project. These two types of evaluations require different data sources. The following table describes the two types of evaluation and their corresponding data sources.

Process evaluation can make use of multiple data sources, both quantitative and qualitative. In contrast, monitoring of results tends to involve the two data sources that produce quantifiable data: service statistics and surveys among members of the intended audience (in particular, people who have participated in project activities in most cases).

***Box 23: Correspondence between
Data Sources and Types of Evaluation***

Data Source	Process Evaluation	Monitoring of Results
Quantitative		
Interviews	X	X
Service statistics	X	X
Qualitative		
Observation		
Direct	X	
Mystery Client	X	
Ethnographic	X ⁷	
Focus Groups	X	

In the following chapter we present a series of illustrative projects. These correspond to illustrative projects or interventions being implemented by NGOs in developing countries (See Box 4.). The tables in Chapter 5 provide guidance on possible indicators and sources of data to readers who recognize their own project from those presented.

Appendix A presents a summary inventory of those indicators most frequently used in evaluating HIV/AIDS prevention projects.

⁷Not recommended for evaluation purposes.

Chapter 5

HIV/AIDS Illustrative Projects

Previous chapters have covered various types of evaluation, indicators, and sources of data. In this chapter, we link these different concepts.

Although HIV/AIDS projects take many different forms in countries around the world, certain illustrative interventions are common. In this chapter, we present 12 illustrative projects, along with a brief description of the intervention, typical objectives, indicators for evaluation process and monitoring results, their respective data sources, and comments on evaluating each type of project.

Many organizations have more than one type of project. In such cases, each requires a separate evaluation.

In Chapter 2 we described types of evaluations. In this chapter we classify the indicators according to the same scheme.

Some readers may question what the difference is between indicators for process evaluation versus monitoring of results for each illustrative project. To summarize the explanation given in Chapter 2, process indicators tend to measure “how much we have done” (e.g., in terms of number of training

events completed, materials produced, events conducted) and “how well we have done it” (with respect to quality of performance and client satisfaction). In contrast, monitoring of results serves to measure achievement of the objectives, most often defined in terms of use of services or changes in KAP (knowledge, attitude, practices).

This distinction between process evaluation and monitoring of results applies to the majority of indicators. However, the line blurs in some cases. Let’s take the example of a VCT program, the objective of which is to increase the number of persons tested for HIV. In this case, the indicator “number of tests completed” might be an appropriate indicator of results. However, one might question whether a program should get credit for “number of referrals made to the VCT program,” given that we have no guarantee that the persons referred will actually use the VCT service. In this manual we have categorized number of referrals as an indicator of results, given that it relates to service use. However, not all readers might agree with this classification.

A second indicator that raises questions regarding classification relates to the distribution and sale of condoms. In this manual, we classify this indicator as a result, given that the purchase of a condom implies probable future use, and increasing the use of condoms is the objective of many HIV/AIDS prevention programs. However, the free distribution of condoms is less clear, since the probability of use is lower. Nonetheless, it seems logical to classify the distribution of condoms — be it through sales or free distribution — in the same category, and in this manual we have considered this indicator to be a result.

The illustrative projects presented in detail below are the same as those listed in Box 4 (Chapter 2) and again in Box 24 following.

***Box 24: List of Illustrative Projects
for NGOs Worldwide***

1. Telephone Hotline on STI/HIV/AIDS
2. School-based Education on STIs and HIV/AIDS for Youth
3. Peer Education
4. National or Regional HIV/AIDS Information Centers
5. Voluntary Counseling and Testing (VCT)
6. Integration of STI/HIV/AIDSs into MCH and Family Planning Programs
7. Prevention of Mother-to-Child Transmission (PMTCT) of HIV
8. Condom Social Marketing
9. Empowerment of Female Commercial Sex Workers (CSWs)
10. Promotion of Condom Use Among High-risk (and Often Predominantly Male) Audiences (e.g., Clients of CSWs, Men Who Have Sex with Men, Prisoners, and Military Personnel)

Box 25: Illustrative Project 1 – Telephone Hotline on STI/ HIV/AIDS

Illustrative Objectives:⁸

- to provide members of the general public with an anonymous means of receiving information and assistance on:
 - modes of STI and HIV/AIDS transmission
 - methods of STI and HIV/AIDS prevention
 - correct use of condoms
 - communication about human sexuality
 - sexual preference
 - human rights
- to promote safer sexual practices to avoid transmission of STIs and HIV (the ABCs):
 - abstinence
 - being faithful
 - condom use (correct and continuous)
- to provide referrals for HIV tests, clinical services, and psychological services

Intervention:

The project establishes a hotline service, publicizing the phone number widely through multiple channels. Counselors are trained to discuss a wide variety of issues with callers (e.g., HIV/AIDS, human sexuality) and to make referrals to available clinical and social services.

Process indicators	Data source
• number of methods used to publicize hotline number	program records
• number of spots broadcast announcing hotline (by channel)	program records, station logs
• % of callers that mention learning about the hotline from the spot	caller survey
• % of callers that report having called before	caller survey
• number of hours service per week	program records
• level of user satisfaction with the quality of the information provided	caller survey

Box 25: Illustrative Project 1 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none">• number of calls by gender, type of question and reason for call	daily logs of callers
<ul style="list-style-type: none">• number of referrals made by type of service (e.g., HIV tests, clinical services, psychological support)	daily logs of callers
<ul style="list-style-type: none">• % of community members aware of the hotline	population-based survey

Comments:

- Given the anonymous nature of this program, it is impossible to do a follow-up survey of callers to measure change in knowledge or attitudes.
- The mystery client technique is a useful way to evaluate the technical quality of the services.
- To assess caller satisfaction, the hotline staff member can ask a few questions at the end of the call. Although courtesy bias is likely, the anonymity of the call counterbalances this tendency.
- In a country where various languages are spoken, the evaluation should indicate the availability of staff to field calls in the most common languages.

⁸The objectives for the illustrative projects are not “SMART.” They were not worded in a “SMART” manner because the exact details would vary from project to project. An example of a SMART initiative would be: to increase the percentage of the general public with anonymous means of receiving information and assistance from 5 percent to 20 percent by 2006.

Box 26: Illustrative Project 2 – School-based Education on STIs and HIV/AIDS for Youth

Illustrative Objectives:

- to increase knowledge of young people on:
 - modes of STI and HIV/AIDS transmission
 - methods of STI and HIV/AIDS prevention (the ABCs)
 - risk factors (e.g., drugs, alcohol) for STI and HIV/AIDS
 - life skills
- to increase self-esteem
- to increase acceptance of people living with HIV/AIDS

Intervention:

The intervention consists of:

- determining the appropriate topics to be covered
- designing a curriculum
- training people to facilitate sessions
- developing educational and informational materials
- implementing project among the target population

Process indicators	Data source
<ul style="list-style-type: none"> • number & characteristics of students that participate in the program by sex, age, grade, and school 	program records
<ul style="list-style-type: none"> • number of hours per week dedicated by school to AIDS prevention activities, by school 	program records
<ul style="list-style-type: none"> • number of materials designed, by topic 	program records
<ul style="list-style-type: none"> • number of materials distributed, by topic 	program records
<ul style="list-style-type: none"> • number of teachers and students trained 	program records
<ul style="list-style-type: none"> • number of schools participating in the project 	program records
<ul style="list-style-type: none"> • level of satisfaction among students and teachers with different aspects of the program 	survey or focus groups with students and teachers

Box 26: Illustrative Project 2 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none">● % of students that know:<ul style="list-style-type: none">– modes of HIV/AIDS and STI transmission– methods of HIV/AIDS and STI prevention (the ABCs)– risk factors for STI and HIV/AIDS– basic life skills	pre- and posttest of KAP (knowledge, attitude, and practice)
<ul style="list-style-type: none">● % of students that increase their level of self-esteem from a score of ___ to ___ in relation to the questions answered in the pre and post test	pre- and posttest survey (KAP)
<ul style="list-style-type: none">● % of students who express acceptance of people living with HIV/AIDS	pre- and posttest survey (KAP)
<p>Comments: Given the confidential nature of the information and the level of education of the participants, it is appropriate to use self-administered interviews. Respondents will be more comfortable and will probably give more candid answers.</p>	

Box 27: Illustrative Project 3 – Peer Education⁹**Illustrative Objectives:**

- to increase knowledge of STI and HIV/AIDS in youth selected to serve as peer educators
- to develop skills of peer educators in:
 - communication techniques
 - ability to refer to other services
 - to increase knowledge of STI and HIV/AIDS among young people through peer educators on:
 - modes of transmission
 - methods of prevention
 - risk factors (drugs, alcohol, etc.)
 - sources of services (condom sales, psychological services, etc.)
- to increase acceptance of people living with HIV/AIDS
- to achieve a multiplier effect, whereby trained youth teach their peers about STI and HIV/AIDS prevention

Intervention:

Project staff work with schools or youth-friendly organizations to select and train a group of young people on STI and HIV/AIDS topics, so they can share this knowledge with other youth in their surroundings. The project can operate within and outside the formal education sector.

Process indicators	Data source
<ul style="list-style-type: none">• number of youth (or other peers) trained as peer educators, by sex, age, place of residence, school/organization, etc.	program records
<ul style="list-style-type: none">• level of participant satisfaction with different aspects of the program	structured or in-depth interviews and/or focus groups

⁹Peer education is not limited to youth; it may involve other “profiles” (e.g., commercial sex workers).

Box 27: Illustrative Project 3 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none">• % of trained youth that know symptoms, modes of transmission and prevention of STI and HIV/AIDS (the ABCs)	pre- and post-test survey (KAP)
<ul style="list-style-type: none">• % of youth that know a source of services (e.g., condom sales, psychological services)	pre- and post-test survey (KAP)
<ul style="list-style-type: none">• number of referrals made to STI and HIV/AIDS services by peer educators	program records
<ul style="list-style-type: none">• % of trained youth with accepting attitudes towards people living with HIV/AIDS	pre- and post-test survey (KAP)
<ul style="list-style-type: none">• number of trained peer educators that carry out educational activities with peers	program records, interviews with peer educators
<ul style="list-style-type: none">• number and type of contacts by peer educators with youth or other peers, per month	program records, interviews with peer educators
Comments: The data for monitoring results are based on a self-report by the peer educators regarding the extent of their educational activities with other youth and, thus, they may be of questionable reliability. Moreover, it is very difficult to assess the extent of knowledge or attitude change among other youth resulting from the efforts of the peer educators.	

Box 28: Illustrative Project 4 – National or Regional HIV/AIDS Information Centers

Illustrative Objectives:

- to increase access to information on HIV/AIDS to non-governmental organizations (NGOs), governmental organizations, and the general public
- to increase the use of information centers by NGOs, government offices, educational establishments, and the general public
- to establish an electronic network for consultations on HIV/AIDS at the local, national, or regional level

Intervention:

Personnel are trained and equipped to operate an information center that serves the needs of staff from NGOs, governmental offices, educational establishments, and others to get information on HIV/AIDS and to use electronic media for information exchange.

Process indicators	Data source
<ul style="list-style-type: none"> • number of media used to publicize the services of the information center 	program records, interviews with staff
<ul style="list-style-type: none"> • number and type of services offered by the center (bibliographical search, use of e-mail) 	program records, interviews with staff
<ul style="list-style-type: none"> • % of users satisfied with different aspects of the center and suggestions for improvement 	structured interview with users

Box 28: Illustrative Project 4 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none">• number of consultations at the information center	log of visits or consultations
<ul style="list-style-type: none">• number of hits on the center's Web site	Web site report
<ul style="list-style-type: none">• number of NGOs affiliated with the Center's listserv	program records
<ul style="list-style-type: none">• number of users of Internet or email services to get HIV/AIDS information, by organization or country of origin	program records
<ul style="list-style-type: none">• % of community members that use the center	population-based survey
<ul style="list-style-type: none">• number and name of the NGOs, government organizations, and educational establishments that visit the center	program records
<p>Comments: Consultations include both in-person visits and contacts made by phone, fax, or email.</p>	

Box 29: Illustrative Project 5 – Voluntary Counseling and Testing (VCT)

Illustrative Objectives:

- to increase the number and percentage of people who get tested for HIV
- to achieve counseling of clients prior to testing in 100% of the cases
- to increase to ___% the number of persons who return for their results and posttest counseling
- to increase the practice of safe sex among those tested

Intervention:

VCT services provide clients with counseling prior to the HIV test on what to expect, when the results will be available, how to protect themselves and others from HIV transmission, and related topics. They obtain consent, draw blood, and subsequently provide the results of the test (either the same day or several days later, depending on available equipment). Clients may receive posttesting counseling, especially those who test positive, and referral to other services, where applicable.

Process indicators	Data source
• number of media used to publicize VCT services	program records
• number of spots broadcast per month to publicize the service	program records, station broadcast logs
• number of hours the service is open per week	program records
• availability of service "after hours" at least one day a week	program records, interview with staff
• availability of functioning equipment to perform the HIV test	facility audit
• availability of staff trained in HIV/AIDS counseling	facility audit
• number of educational materials distributed	program records
• level of satisfaction with services among clients	client exit interview

Box 29: Illustrative Project 5 (cont.)

Result/change indicators	Data source
• number of clients attending pre-counseling, by sex, age, marital status, and place of origin	program records
• number of HIV tests performed per month	program records
• number of condoms distributed	program records
• % of clients that receive counseling before the HIV test	program records
• % of those tested that return for their results	program records
• % of clients tested that receive counseling after they receive their results	program records
• profile of clients that do not return for their results (e.g., socio-demographic and other characteristics)	program records
• % of the population tested for HIV	population-based survey
• % of community members aware of VCT services	population-based survey
Comments Given the sensitivity of VCT, the evaluator should use methods that protect the confidentiality of the clients. He/she should collect as much data as possible at the time of the initial visit to the VCT service (in those cases where the client must return for the results), given the difficulties of collecting information subsequently.	

Box 30: Illustrative Project 6 – Integration of STI/HIV/AIDS into MCH and Family Planning Programs

Illustrative Objectives:

- to increase knowledge of MCH and FP clients on:
 - modes of STI, HIV/AIDS transmission
 - methods of STI, HIV/AIDS prevention (the ABCs)
 - correct use of condoms
 - places to go for information, diagnosis, and treatment of STIs and HIV/AIDS
- to increase communication among couples on topics related to sexual behavior and HIV transmission
- to increase the use of safer sexual practices to avoid sexual transmission of HIV, through:
 - mutual fidelity
 - correct and consistent condom use in sexual relations
- to increase the number of referrals to other services (VCT, STI diagnosis and treatment)

Intervention:

Project staff train MCH and FP service providers (including clinical personnel, educators, and counselors) in STI and HIV/AIDS, so that they can integrate these topics into the routine delivery of MCH and FP services (e.g., in group talks or individual counseling sessions)

Process indicators	Data source
<ul style="list-style-type: none"> • number of educational talks about STIs and HIV/AIDS per month 	program records
<ul style="list-style-type: none"> • % of clients in MCH/FP services that receive counseling on STIs or HIV/AIDS through educational talks or individual counseling 	program records
<ul style="list-style-type: none"> • level of client satisfaction with different aspects of the counseling 	client exit interview, focus groups

Box 30: Illustrative Project 6 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none"> • % of MCH/FP clients that know: <ul style="list-style-type: none"> – 3 modes of STI and HIV/AIDS transmission – 3 methods of STI and HIV/AIDS prevention – how to recognize a STI – the role of STIs in the transmission of HIV 	client exit interview
<ul style="list-style-type: none"> • number of referrals to other services by type of service (VCT, STI diagnosis, and treatment) per month 	program records
<ul style="list-style-type: none"> • % of FP/MCH clients that know at least one source of VCT and STI services 	client exit interview
<p>Comments:</p> <ul style="list-style-type: none"> • The evaluator can measure the level of knowledge among MCH/FP clients after counseling using a client exit interview. However, in the absence of a “pre-counseling” measure, it is not possible to evaluate change. (As an alternative, the evaluator could compare the knowledge of this group of clients with a control or comparison group in which MCH/FP staff did not receive the training). • Programs often encourage husband-wife communication on sexual matters as an important first step to the practice of safer sex. However, it is difficult to evaluate this objective, since the evaluator loses contact with the client after she leaves the clinic. • Mystery (simulated) clients can be useful in assessing the technical quality of the counseling. 	

Box 31: Illustrative Project 7 – Prevention of Mother-to-Child Transmission (PMTCT) of HIV

Illustrative Objectives:

- to increase access of pregnant women to HIV counseling and testing services
- to increase access of HIV+ pregnant women to care and support services and anti-retroviral therapy (ARVs) that reduces the risk of MTCT

Intervention:

Many PMTCT interventions involve just the provision of ARVs to HIV+ pregnant women. The current trend toward PMTCT+ looks at a broader range of needs and services, not just ARV access (e.g., addressing issues related to care and support of HIV+ women, increasing women’s access to family planning to prevent unintended pregnancies, and promoting HIV risk-reduction practices among all women of reproductive age). However, in an effort to simplify our M&E discussion, we will assume that our illustrative PMTCT project is limited to the activities corresponding to the objectives listed above, namely increasing VCT access among pregnant women and increasing ARV and care/support access for HIV+ women. We will also assume that the project addresses issues of both “supply” (e.g., service availability) and demand (e.g., promoting service utilization).

Process indicators	Data source
<ul style="list-style-type: none"> • number of antenatal care (ANC) sites offering PMTCT+ services 	program records
<ul style="list-style-type: none"> • % of health care workers newly trained or re-trained in the minimum PMTCT package 	program records, training logs
<ul style="list-style-type: none"> • number of antenatal care sites with no stockouts of rapid HIV test kits in the last 12 months 	program records
<ul style="list-style-type: none"> • number of radio spots aired promoting ANC use in the past 12 months 	program records
<ul style="list-style-type: none"> • number of PMTCT/PMTCT+ sites that report no stockouts of ARVs in the last 12 months 	program records

Box 31: Illustrative Project 7 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none"> • % of pregnant women attending ANC in the last 12 months who were: (a) offered an HIV test; (b) accepted the test; and (c) received posttest counseling and HIV results 	KAP survey, exit interviews
<ul style="list-style-type: none"> • % of HIV+ women who delivered in the past 12 months and received a complete course of ARVs for PMTCT 	program records
<ul style="list-style-type: none"> • % of the target population that knows the three ways HIV can be transmitted from mother to baby 	pre- and posttest survey (KAP)
<ul style="list-style-type: none"> • % of HIV+ pregnant women identified through counseling and testing in the last 12 months who were referred for care and support services 	program records, exit interviews
<ul style="list-style-type: none"> • % of HIV+ women who were counseled on infant feeding options 	program records, exit interviews
<ul style="list-style-type: none"> • % of women who delivered in the past 12 months who received postpartum family planning counseling and/or services 	KAP survey
<p>Comments:</p> <ul style="list-style-type: none"> • M&E activities related to PMTCT will usually rely on a broad range of methods and data sources. • The ability to document success in the area of PMTCT is very dependent upon the degree to which: (a) women come in contact with the formal health system and (b) there is adequate follow-up of HIV+ women to track key processes and outcomes. • The stigma associated with HIV or confidentiality issues around disclosing an individual's HIV status might make it difficult to identify HIV+ women. 	

***Box 32: Illustrative Project 8 –
Condom Social Marketing***

Illustrative Objectives:

- to increase knowledge of condoms as a method of HIV prevention
- to increase condom use among sexually active adults (frequently in a specific target population)

Intervention:

The primary activity is the promotion and sale of condoms at a favorable price for the target population (often through government subsidiaries or international donor agencies). In recent years social marketing has also been used to promote VCT. The social marketing approach draws on commercial marketing in a number of ways: careful evaluation of consumer preferences in terms of brand names and packaging; positioning the product for the specific target population (e.g., a certain socioeconomic class, men who have sex with men); aggressive promotion of products through mass media (radio, TV, billboards, posters, T-shirts); and point-of-purchase advertising.

Process indicators	Data source
<ul style="list-style-type: none"> • number of spots produced 	program records
<ul style="list-style-type: none"> • number of spots broadcast, by channel (radio, TV) 	program records, station broadcast log
<ul style="list-style-type: none"> • number of items distributed by product type (e.g., posters, T-shirts) 	program records
<ul style="list-style-type: none"> • % of target population that has seen or heard a specific message 	special (preferably representative) survey of the target population
<ul style="list-style-type: none"> • % of population that has seen or heard any of the campaign messages through any of the channels 	special (preferably representative) survey of target population

Box 32: Illustrative Project 8 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none"> • number of condoms sold per month 	program records
<ul style="list-style-type: none"> • % of target population that recognizes condoms as a method of preventing HIV transmission 	special (preferably representative) survey of target population
<ul style="list-style-type: none"> • % of target population that has heard the name of the social marketing brand of condom 	special (preferably representative) survey of target population
<ul style="list-style-type: none"> • % of the target population that has bought the specific brand of condoms during a determined period of time (e.g., the last month) 	special (preferably representative) survey of target population

Comments:

- Generally the key indicator of results is the number of condoms sold, given the ease of obtaining this information and the commercial nature of the program. The volume of sales is available from the start and provides the primary indicator of results. Special surveys can be implemented to obtain more data, but these should be conducted less frequently, because (1) they are costly and (2) they require relatively large samples to detect significant change in purchase behavior of condoms.
- In social marketing projects much of the research/evaluation budget goes to the design and promotion of the product (e.g., test of brands and packaging, diagnostic research in target populations, and test messages).
- Increasingly, social marketing programs promote the ABCs rather than just C (condom use). In such cases, the surveys used to evaluate change would include KAP items related to abstinence, monogamy, and partner reduction.

Box 33: Illustrative Project 9 – Empowerment of Female Commercial Sex Workers (CSWs)

Illustrative Objectives

- to increase knowledge of:
 - STIs, HIV/AIDS,
 - sources of information and services
- to increase comprehensive medical care for CSWs (obstetrical/gynecological, diagnosis, treatment of STIs)
- to increase access of CSWs to HIV testing at an affordable price
- to increase condom use in sexual relations with clients and partners

Intervention:

Although the types of activities conducted vary by country, this type of project often provides a safe haven for CSWs, which they can use throughout the day to receive a variety of services (e.g., medical, psychological, dental, beauty classes, educational talks on STIs and HIV/AIDS, condoms, informational pamphlets). Some projects also provide other social activities (e.g., exercise classes, cooking facilities on the premises, magazines to read, coffee). This safe haven with the above-mentioned services is designed to improve CSW self-esteem and give them the tools to protect themselves from HIV. One example of this type of project is “La Sala”, which operates in Costa Rica and Guatemala, and similar programs in neighboring countries.

Process indicators	Data source
<ul style="list-style-type: none"> • number of CSWs by age, origin, marital status, place participating in the program 	<p>program records</p>
<ul style="list-style-type: none"> • level of satisfaction with different aspects of the program 	<p>structured or in-depth interviews and/or focus groups</p>

Box 33: Illustrative Project 9 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none">• number of CSWs that receive medical consultations	program statistics
<ul style="list-style-type: none">• number of CSWs referred to other services (e.g., psychological, diagnosis, and treatment of STI, dental)	program statistics
<ul style="list-style-type: none">• number of CSWs that receive (or are referred for) VCT	program statistics
<ul style="list-style-type: none">• % of CSWs that know forms of prevention and transmission of STI and HIV/AIDS	surveys with CSWs
<ul style="list-style-type: none">• % of CSWs that know the most common STDs (gonorrhea, syphilis, herpes, hepatitis B)	surveys with CSWs
<ul style="list-style-type: none">• % of CSWs that know a place where they can get an HIV test	surveys with CSWs
<ul style="list-style-type: none">• % of CSWs that report “always” using a condom with all their clients	surveys with CSWs
<ul style="list-style-type: none">• % of CSWs that report having used a condom at last sexual relations	surveys with CSWs
Comments: Evaluators may have difficulty gaining access to this group, given their work environment.	

Box 34: Illustrative Project 10 – Promotion of Condom Use Among High-risk (and Often Predominantly Male) Audiences (Clients of CSWs, Men Who Have Sex with Men, Prisoners, Military Personnel)

Illustrative Objectives:

- to increase knowledge among intended audience of:
 - modes of HIV/AIDS and STI transmission
 - methods of HIV/AIDS and STI prevention
 - correct and consistent use of condoms
 - source of information about STIs and HIV/AIDS
- to increase correct condom use (using an anatomical model)
- to increase condom use in each sexual relation

Intervention:

Trained personnel arrange to visit locations where the intended audience tends to congregate, and they give education talks, covering the basic facts about STIs and HIV/AIDS, and correct condom use. Also, they often distribute condoms, and they may demonstrate the correct use of the condom using an anatomical model. In addition to condoms, they may discuss the benefits of monogamy or at least reducing the number of partners.

Process indicators	Data source
<ul style="list-style-type: none"> • number of people reached by sex, age, type of group per month 	program records
<ul style="list-style-type: none"> • number of geographical areas covered each month 	program records
<ul style="list-style-type: none"> • level of satisfaction of participants with information given 	exit or indepth interviews and/or focus groups

Box 34: Illustrative Project 10 (cont.)

Result/change indicators	Data source
<ul style="list-style-type: none">• number of condoms distributed per month	program records
<ul style="list-style-type: none">• % of participants that know methods of transmission and prevention of STI and HIV/AIDS	pre- and posttest survey (KAP), exit interviews
<ul style="list-style-type: none">• % of participants that know one source for:<ul style="list-style-type: none">– VCT– diagnosis and treatment of STIs– condoms	pre- and posttest survey (KAP), exit interviews
<ul style="list-style-type: none">• % of participants that correctly place a condom on an anatomical model	observation (during exit interview)
<ul style="list-style-type: none">• % of participants that report having used a condom in their last sexual encounter	exit interviews
Comments: Given the nature of the intervention, any questionnaire used should be short, simple, and easy to apply.	

Chapter 6

Data Processing and Analysis

Once the data are collected, the next steps in the process are editing, coding, cleaning, processing, and analyzing it. We'll begin with quantitative data, then move to qualitative data.

Quantitative Data

After data collection, the evaluator should organize the information carefully, handling the instruments from different types of data collection (hotline surveys, client exit interviews) separately. Each case should have a unique identification number.

Editing

This process consists of verifying that all the instruments have been filled out correctly. Field supervisors carry out part of this process in the field, but research personnel back in the office are responsible for subsequent revision.

Specifically, they review the responses on the questionnaires to ensure that they are legible and logical. To the extent possible and without making up responses, they try to clarify any

questionable entries, consulting if possible with the interviewer who did the interview. The evaluator also identifies and trains the staff to look for logic in the responses. For example, is it possible for a 20-year-old woman to have five children? In short, the evaluator wants the information to be “clean” before proceeding to the data entry process. If data are incomplete and can not be recovered from the primary source, the evaluator should acknowledge the problem of missing data in the final analysis.

Coding

Surveys generate an enormous amount of information. Coding is a method by which the verbal responses of those interviewed are transformed into manageable form for subsequent processing and analysis (generally by computer, but at times, by hand).

Coding consists of transforming responses into codes to be used in processing and analyzing data. Many variables are pre-coded (for example: 1 = male and 2 = female). In other cases, the response to the question is the “code” (example: “How old are you?” = 29). However, for open-ended questions (example: “Why do you think you are at risk of contracting AIDS?”), it is necessary to create and assign a unique code to each response or group of highly similar responses.

One approach to coding process is as follows:

- Review 10% of the questionnaires collected and make a list of most common responses to open-ended questions.
- Create categories that capture each different answer or group of similar answers. For example: “Craftsmen” includes shoemakers, cabinetmakers, and ceramic or clay sculptors.

Data Source

- Prepare a list of codes, assigning a number to each category.¹⁰
- Develop a list of new codes that emerge during the coding process, and assign a number to each new response.
- Review the codes for each variable to ensure all responses have an assigned code.

Creating a Database

A database is a computerized archive that stores information. In creating a database, the evaluator defines the position of all the variables included.

Various programs for creating databases exist on the market. EPI INFO, ACCESS, and SPSS are frequently used to manage quantitative information. These programs facilitate data management, are easy to use and, in the case of EPI INFO, require a relatively small amount of memory on the computer.

Data Entry

This process consists of entering all of the information collected on the instruments into a computerized database that permits the quick and optimal management of information.

We recommend using or hiring a person with data entry experience (even if this requires paying for the service outside of the organization) for several reasons:

- greater accuracy;

¹⁰ Coding with numbers is used more frequently than coding with letters. Some computing programs can combine numbers and letters (alphanumeric).

-
-
- greater speed of data entry; and
 - lower cost (when done correctly the first time).

The larger the volume of data, the more important that the entry be done accurately, promptly, and economically. Even though the organization may hire an external consultant, it is important to build local and organizational capacity in this area. In addition to having data management expertise, ideally the person will also be willing to mentor local staff.

Data Cleaning

This process ensures that the data used for analysis contain the least possible number of errors.

Some programs, such as EPI INFO, ACCESS, and SPSS, have a mechanism that allows the evaluator to define what codes are acceptable on a given variable to be entered and then rejects all other values for that variable during data entry.

For example: Let's assume that the variable "marital status" has only four acceptable values: 1 = married, 2 = consensual union, 3 = single, and 8 = other. The program allows only these data codes to be entered and thus reduces errors.

The evaluator can also introduce a test of logic. He/she defines situations that are not logical, and the program attempts to identify them in the database (e.g., the woman under 20 years old with five children). Sometimes a review of the original questionnaire shows that this implausible case did in fact occur. In others, the error becomes apparent. Cleaning the data helps correct many of the errors before the analysis begins.

Data Analysis

In the analysis of quantitative data, the evaluator usually begins

by obtaining frequencies of all the variables in the data set. Frequencies tell us, in absolute numbers or percentages, how many times a given response occurs. This step provides a general idea of the results, as well as possible errors that still exist in the data.

Frequencies provide useful results on the majority of the process indicators. However, evaluators are often interested in analyzing one variable in relation to a second factor: for example, the level of satisfaction with services, by clinic. To illustrate, we will use the example of the survey on client satisfaction with VCT services for two different clinics (See Box 35.). Suppose we have 300 clients in the first clinic, A, and 200 clients in the second, B.

Box 35: Analyzing Variables		
Level of Satisfaction	Clinic A	Clinic B
Excellent – Good	80% (240)	90% (180)
Fair – Poor	20% (60)	10% (20)
	n = (300)	n = (200)
*Based on the question: "Overall, how would you rate the service?"		
	$\chi^2 = 8.93$	$p = 0.002$

In this example, it appears that the clients from Clinic B were more satisfied than the clients from Clinic A. To arrive at this conclusion, we must determine if the difference in percent

satisfied (80% in A, 90% in B) is significant, or if it is so small that it could have occurred by chance.

We apply the statistical technique known as chi square to determine if the observed difference is statistically significant. Chi square (χ^2) is interpreted using the p-value (from statistical analysis programs such as Epi Info or SPSS). If the p-value is less than 0.05, we can conclude that the difference is significant. On the other hand, if the p-value is greater than 0.05, it is not considered significant.¹¹

We also use cross tabulations in measuring change. Let's return to the example of the evaluation of results described in Box 5 (the project to increase correct condom use among commercial sex workers). The hypothesis of this evaluation is that correct condom use will increase after the educational workshop. Let's assume that we have a total of 400 CSWs. The dependent variable (the outcome to be explained) is "correct condom use." The independent (or explanatory) variable is "participation in

Box 36: Has Correct Knowledge of Three Steps in Correct Condom Use

	Workshop Participant	Non-workshop Participant	Total
Has Knowledge	80% (160)	35% (70)	(230)
Does Not Have Knowledge	20% (40)	65% (130)	(170)
	(200)	(200)	(400)

¹¹ For more information on the statistical interpretations, readers should consult a statistician or a statistics textbook.

the workshop.” This cross tabulation measures whether participation in the workshop is correlated with knowledge of correct condom use.

In this example, we observe that the percentage of CSWs who could correctly place a condom on the anatomical model was higher for those who participated in the workshop. The p-value is less than 0.05.

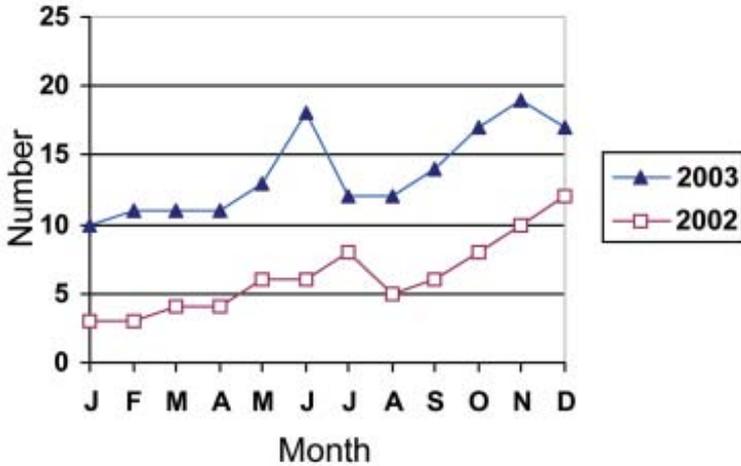
Illustrative Presentation of Quantitative Results

Evaluators often present the results of an evaluation in the form of a written report. Frequently, they will also prepare the results in graphic form as part of a PowerPoint presentation. On the pages that follow, we present illustrative results from an evaluation of Illustrative Project 5 (VCT). Graphs of this sort can accompany the text of the final report or serve as the basis of an oral presentation.

***Box 37: Voluntary Counseling and Testing (VCT) –
Promotion and Availability of the Service
(3 Clinics)***

Indicator	2001	2002
* number of media used to publicize the VCT service	1	2
* number of hours of service per week	40 hours/ week	40 hours/ week
* number of hours of service "after hours"	0	3 hours/ week

Box 38: Number of Spots Broadcast per Month, by Year (All Media Combined)



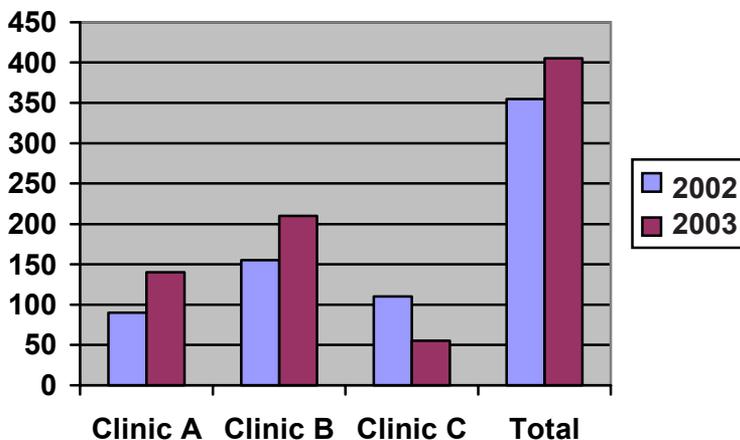
Box 39: Availability of Necessary Equipment in Good Working Order to Perform the HIV Test

	Clinic A		Clinic B		Clinic C	
	2002	2003	2002	2003	2002	2003
Equipment required:						
• sterile gloves	✓	✓	✓	✓	✓	✓
• syringes		✓	✓	✓	✓	✓
• chlorine solution	✓	✓	✓	✓		
• disposal for contaminated items	✓	✓	✓	✓	✓	✓
Conditions:						
• waiting room clean	✓	✓	✓	✓		
• sufficient seating for clients	✓	✓	✓	✓		

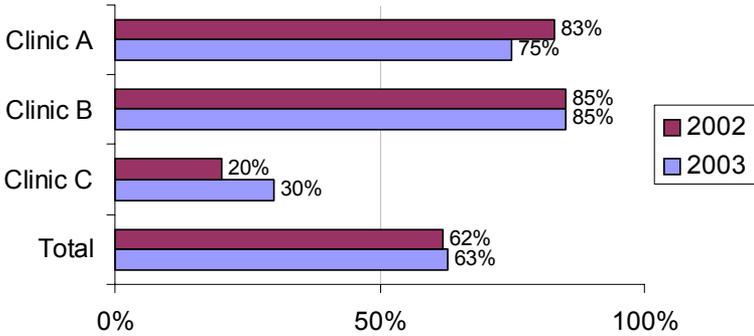
***Box 40: Availability of Personnel
Trained in VCT Counseling***

	Clinic A	Clinic B	Clinic C	Total
Number who give counseling				
2002	2	2	3	7
2003	2	2	3	7
Of those, number trained in counseling				
2002	1	0	0	1
2003	2	1	0	3

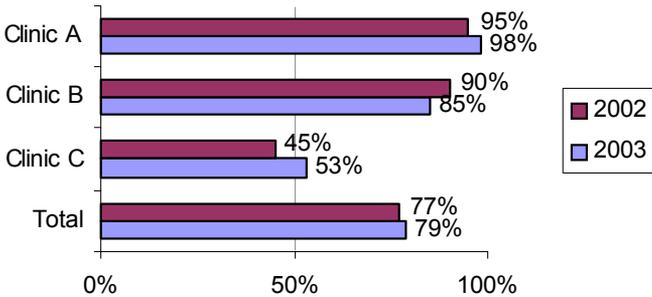
***Box 41: Distribution of Materials –
Number of Pamphlets Distributed***



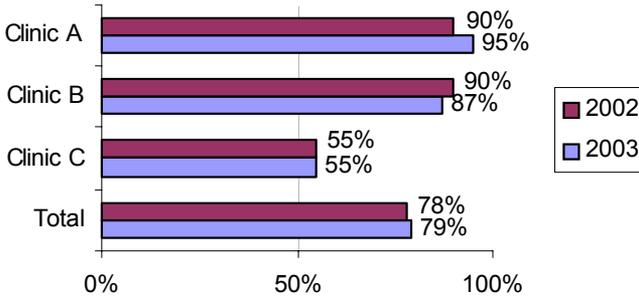
**Box 42: Level of Satisfaction with Services –
(a) The waiting time was acceptable.**



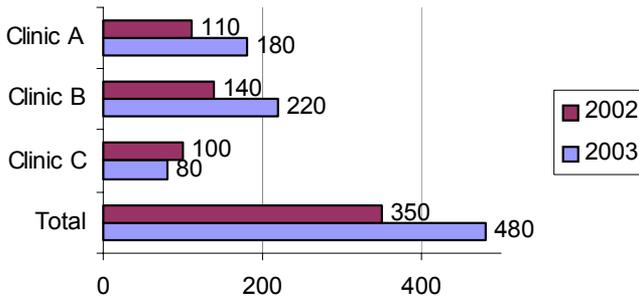
**Box 43: Level of Satisfaction with Services –
(b) Treated with respect by staff**



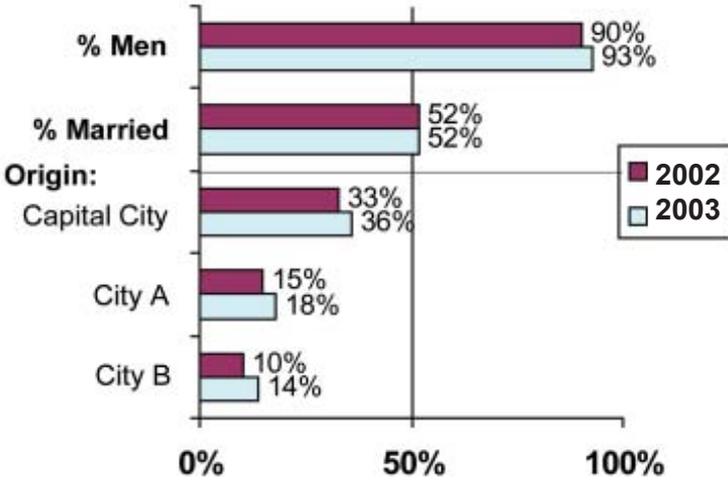
**Box 44: Level of Satisfaction with Services –
(c) Felt comfortable asking questions**



Box 45: Number of Users Attending Precounseling



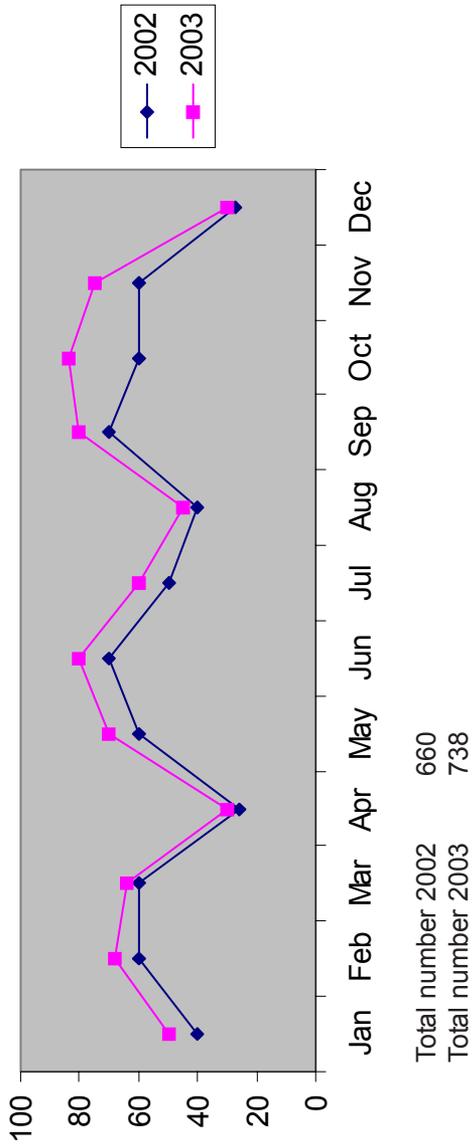
***Box 46: Characteristics of the Users
(All Clinics Combined)***



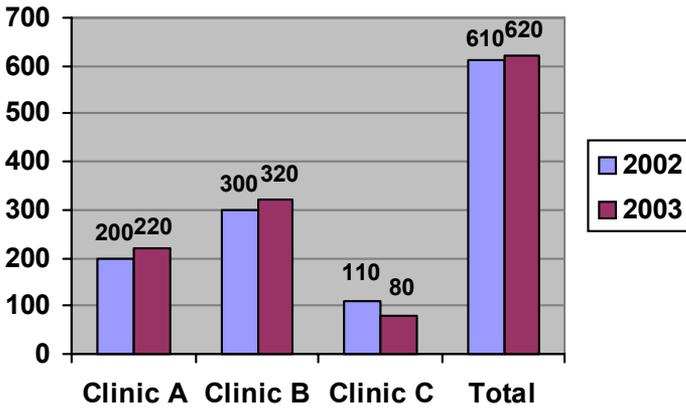
***Box 47: Characteristics of the Users
(All Clinics Combined)***

	Clinic A		Clinic B		Clinic C		Total	
	2002	2003	2002	2003	2002	2003	2002	2003
Gender: % men	95%	95%	87%	92%	88%	92%	90%	93%
Age: average (years)	38.0	37.0	36.0	36.0	39.0	37.0	37.7	36.7
Marital Status: % married	55%	55%	50%	51%	51%	50%	52%	52%
Origin								
• Capital	37%	40%	32%	33%	30%	35%	33%	36%
• City A	12%	13%	17%	19%	16%	22%	15%	18%
• City B	9%	11%	13%	17%	8%	14%	10%	14%

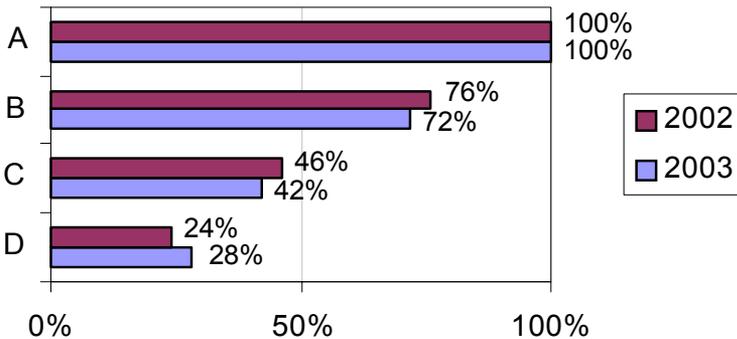
**Box 48: Number of Tests Carried Out by Month
(3 Clinics)**



**Box 49: Number of Condoms Distributed
(by Clinic)**



**Box 50: Counseling Before and After the HIV Test
(3 Clinics Combined)**



Percentage of users who:

A: received counseling before the HIV test

B: came back for the results of the HIV test

C: came back for the results and received counseling after the test

D: did not come back for the results

Qualitative Data

Processing and analysis differ markedly for quantitative versus qualitative data. Processing quantitative data requires knowledge of statistical procedures and computer operations (if used for analysis). However, the steps are well established. By contrast, qualitative analysis requires a systematic approach but an element of individual judgment. The evaluator must strive to remain highly objective to avoid biasing the results with personal opinions. Because qualitative analysis is not based on numbers (e.g., percentages, averages), other strategies must be used to find patterns in the results.

Various programs exist on the market for analyzing qualitative information, including ANTHROPAC, NUD*IST and ETHNOGRAPH.¹² These programs facilitate the organization of information and help evaluators identify sections of the discussion relevant to a particular topic, points of agreement and disagreement within the groups, and the topics most discussed in the groups.

In this section on how to manage and analyze data once it is collected, we focus on data from focus groups and in-depth interviews.

Preparing the Transcriptions

The evaluation team should:

- Record the in-depth interviews and focus groups on audiocassette, if possible. Identify each tape with the characteristics of the participants (e.g., women 15-24 years old, users or non-users, date and time of the session,

¹² Computerized programs are very useful when the evaluator is managing a large volume of information, for example 20 or more in-depth interviews or focus groups with distinct sub-populations.

name of community). It is useful to assign each focus group or interview a unique identification number.

- Transcribe the information from the tapes into a computerized text file, trying to reproduce verbatim what was recorded on the cassette. If there are sections on the cassette that are inaudible, write “inaudible” but continue with the verbatim transcription once the words become audible. Identify each transcript with the same identification numbers and descriptions used on the cassettes.
- Ensure that the person who carries out the focus group (moderator or observer) or the in-depth interview (interviewer) does the corresponding transcription as soon as possible, preferably before carrying out a new focus group or interview. If the person’s memory is fresh, he or she may be able to reconstruct the sections that are barely audible on the recording. Furthermore, the work will be less laborious, since it is easier to transcribe the cassettes one by one, instead of several at a time.

Processing Data for Analysis

- Review the objectives of the evaluation as a reminder of the key issues to be examined in the analysis.
- Organize the analysis of results according to the study objectives (which often correspond closely to the questions on the discussion guide); that is, make a list of the topics to be covered in the analysis of the tapes.
- If a topic is too broad, try to subdivide it to make the task more manageable.
- Read each transcription, identify the segments (comments) that apply to each topic covered. They will not necessarily be in order. For example, a comment related to the first topic may appear at the end of the transcription. Note: this step can be carried out by hand or by computer (with a qualitative data analysis program).

-
-
- Once all of the comments are categorized by different topics (objectives of the evaluation) and the irrelevant comments are eliminated, review all the text relevant to the first topic and determine the trends in the participants' or interviewers' opinions (for example, almost all in favor, almost all against, or the group is clearly divided in their opinion about the question). Continue to do the same with the remaining topics.
 - If the evaluation included members of different groups (men, women; youth, parents; urban, rural), try to determine if there are differences of opinion among the groups.

Preparing the Report

- Prepare a description of the socio-demographic characteristics of the evaluation participants (e.g., gender, ages, urban/ rural residences, level of education); for this table it is appropriate to use exact numbers (absolute numbers, percentages, averages, etc.).
- For each objective (topic covered), summarize the primary ideas expressed.
- Try to indicate the intensity of opinion on different topics.
- Use verbatim quotes from the transcriptions to summarize the principal opinions or to illustrate key points.
- Avoid presenting the results in quantitative terms (e.g., percentages, averages, absolute numbers). More importantly, present the principal trends using expressions such as:
 - “Most of the participants were in favor of . . .”
 - “No one was opposed to the idea that . . .”
 - “The participants were totally divided in their opinions with respect to . . .”

-
-
- “Almost all of the participants spontaneously mentioned that . . .”

Conclusions and Recommendations

- After presenting the findings, include a section that explains the programmatic implications of the results. Which aspects of the program/project functioned very well and did not need any type of modification? Based on the results, what are the concrete changes suggested?

Using Evaluation Results to Improve Programs

The evaluation only serves its purpose if the results are used in some way. A variety of important audiences exist to whom evaluation findings should be presented.

Audiences for Evaluation Results

Program Administrators or Managers

A good program administrator or manager wants and needs to know how well the organization’s activities are functioning. The evaluation provides concrete indicators of how successful the projects of an organization are at reaching their objectives. If they are achieving them, the evaluation is a powerful tool to solicit future funding and/or maintain financial support. Moreover, the findings can be used to improve the institutional capacity of the organization. If the evaluation shows certain activities are declining, the program managers or administrators have concrete information with which they can make changes midcourse in the project.

Service Providers and Program Staff

We tend to think of program managers as the primary audience for the results of an evaluation. However, the findings are also important to those who actually provide the services and implement the project activities.

First, it is important for service providers and program staff to know what the clients like and dislike about the services or activities offered by the project.

Second, the evaluation indicates that the organization is trying to respond to the needs of the specific sub-populations within the general population, and of the successful work of the organization in completing its mission.

Third, the evaluation sends an implicit message to the intermediate level staff: their work is sufficiently important to merit systematic evaluation. Even though some fear evaluation, in reality, it can be used to help the staff understand that their work is important too. Moreover, in some cases the evaluation results can be used to recognize staff performance (including public acknowledgement from the director of the organization or by some other high level individual), which can improve morale among staff.

Donor Agency Staff

Donors have different options for investing their funds. Donor agency staff are often deeply committed to the programs they fund, and thus evaluation is important to them for a number of reasons. First, they can use the evaluation results to justify continued or increased funding within their own agencies for programs that show results. Second, the evaluation is an indication that the organization is committed to doing the job well in a very systematic way. Even if the results of a particular project are not as positive as hoped, the donors will tend to

view the organization's commitment to evaluation positively. Finally, the presentation of evaluation results in a public forum focuses attention on the area or topic of interest and creates a certain enthusiasm for continuing work in this area. When the evaluation demonstrates to donors that the desired changes have occurred, they may renew their commitment to make future programming in this area.

Program Participants and Community Members

Often evaluators collect information from those participating in the activities of a given project or from those who live in a determined area where the project is operating. However, too little effort has been made to feed information back to these individuals or communities. In many cases, it would not be practical to give this type of feedback (for example, it would be logistically difficult, not to mention a violation of confidentiality, to follow up with individuals who attended a VCT service). However, other evaluations provide results that are easier to share with those participating in the project and involved in the evaluation.

The presentation of data should be in a different format for community groups than more professional audiences. Sharing information with clients, participants, and community members sends a clear message that their opinion matters.

Some Formats for the Presentation of Results

Although an evaluation report is the most common way of presenting the results, it is not the only form of providing this information. Some other available options are:

PowerPoint Presentation

This format is commonly used to highlight the principal points of the evaluation for different audiences. This form of

presentation is particularly effective for those who have neither the time nor the inclination to read a complete report on the evaluation findings. A verbal presentation aided by PowerPoint can be an excellent vehicle to communicate data and stimulate discussion among members of the board of directors, local government officials, NGO staff, donors, and others. It can also be useful in the context of workshops (mentioned below).

Workshops of Results Analysis and Identification of Actions to Implement Improvements

One method to promote the use of evaluation findings is to plan a meeting with people directly involved in the program (e.g., program managers, service providers, and field staff). The key results can form the object of a meeting in which participants examine and discuss the results in detail. It is important to participants in the meeting to recognize that the reason for evaluation is to act on the results.

Reports and Publications

Even if one does a PowerPoint presentation or conducts a workshop, it is important to have a complete report of the evaluation findings.

Specifically, it is useful to document the findings of all the variables or topics included in the evaluation, not only those of great importance. Despite the fact that a limited number of people may read the report, it serves as important documentation of the evaluation process and the results. The evaluators should write these reports in a user-friendly format, avoiding academic language, and they should include graphs to facilitate understanding.

Even though many evaluations are not sufficiently rigorous to be published, some results merit dissemination beyond the organization responsible for the program. The results of an

innovative project that has good evaluation data can be very useful in informing others of the effectiveness of the approach. Moreover, publication brings visibility to the organization in question, and it contributes to advancing knowledge in the area of HIV/AIDS prevention.

Appendix A

An Inventory of Common Indicators¹³

Process Indicators

“How much have we done? (during a specific period of time: month, trimester, year)”

Number of:

- events carried out (educational talks, workshops, support groups, etc.)
- participants in the event or activity (clients, students, sex workers, etc.)
- products or materials designed/produced
- each product (or material) disseminated or distributed
- organizations participating in the activity
- channels used to disseminate the messages
- messages disseminated by channel
- geographic areas covered by the activities

¹³ The indicators listed in this appendix constitute a summary of those presented project by project in Chapter 5.

-
-
- services offered by the program
 - service hours per week
 - hours of service provided after-hours

Availability of:

- equipment necessary to offer the service
- staff necessary to offer the service

“How well have we done it? (in terms of carrying out the activity, not the results obtained)”

Level of satisfaction with the service or activity, measured by multiple aspects:

- the information was understandable
- the information was sufficient
- the presentation was interesting
- the audience could participate (ask questions, etc.)
- staff treated clients/participants with respect

Percentage of the target population that:

- remember having heard or seen the message (e.g., social marketing campaign)
- remember the content of the message

Result/Change Indicators

Knowledge

Percentage of the target population who know:

- three modes of HIV transmission

-
-
- three methods of prevention of HIV and STD transmission (the ABCs)
 - that a condom is a method of preventing HIV transmission
 - a source of condom distribution (alternatives: number of sources known)
 - three actions of correct condom use
 - risk factors for HIV infection

Percentage of people that know a source for:

- voluntary counseling and testing (VCT)
- STD testing
- STD diagnosis and treatment
- condoms

Percentage of people that know:

- the role of STDs in HIV transmission
- % of commercial sex workers that know the most common STDs: gonorrhea, syphilis, herpes, hepatitis B (in programs for sex workers)

Attitudes

% of people that demonstrate their acceptance of people living with HIV/AIDS (PLHA) through the following indicators:

- Would you be willing to share a meal with a person you knew had HIV or AIDS?
- Would you be willing to care for a relative who became sick with the virus in your own household?
- Believe that a female teacher who has the AIDS virus should be allowed to continue teaching in the school?

Skills

- % of women who can negotiate the use of a condom in sexual relations with their husband
- % of people who correctly place the condom on an anatomical model (dildo)

Behaviors or Practices

- % of youth who have abstained from sex
- % of youth/adults who have had only one sexual partner in past 12 months
- number of sexual partners in past 12 month
- % that report having used a condom at last (penetrative) sexual relations
- % that report having used a condom in their last sexual relation with an occasional partner (non-cohabiting)
- % that report having used a condom at last sexual relations (in the case of commercial sex workers: with their client or with their partner)
- % that report “always” using condoms in their sexual relations (or in sexual relations with casual partners)

Other Indicators according to Project Type¹⁴

- number of visits or consultations to the service
- number of condoms distributed

¹⁴Numbers in parenthesis on the following pages refer to the type of project (listed in Box 4).

-
-
- number of people referred to the activity by a specific program announcement

(1) Telephone Hotline on STI/HIV/AIDS

- number of calls, by user characteristics

(3) Peer Education

- number and percentage of youth trained as peer educators
- number and type of contacts with youth (or others)

(4) National or Regional Level HIV/AIDS Information Centers

Number of:

- consultations at the Information Center
- hits or connections to the Center's Web site
- NGOs affiliated with or listed on the Internet
- users that utilize the Internet services or e-mail to obtain information on HIV/AIDS according to origin
- and name of NGOs, GOs, and establishments that visit the Center

(5) Voluntary Counseling and Testing (VCT)

Number of:

- clients, by key characteristics
- tests carried out

Percentage of clients:

-
-
- tested who receive counseling **before** the test
 - tested who receive counseling **after** the test
 - who return for their results after being tested
 - who do not return for their results after being tested, by characteristics
 - profile of clients who do not return for their results after being tested

(7) Prevention of Mother-to-child Transmission

Number of:

- antenatal care sites with no stockouts of rapid HIB test kits in the last 12 months
- number of antenatal care (ANC) sites offering PMTCT+ services
- radio spots aired promoting NAC utilization in the past 12 months
- PMTCT/PMTCT+ sites that report no stockouts of ARVs in the last 12 months

Percentage of:

- health care workers newly trained or retrained in the minimum PMTCT package
- pregnant women attending ANC in the past 12 months
- HIV+ women who delivered in the past 12 months and received a complete course of ARVs for MTCT
- the target population that knows the three ways that HIV can be transmitted from a mother to her baby

-
-
- HIV+ pregnant women identified through counseling and testing in the last 12 months who were referred for care and support services
 - HIV+ women who were counseled on infant feeding options
 - % of women who delivered in the past 12 months who received postpartum family planning counseling and/or services

(8) Condom Social Marketing

- % of the target population that has bought a specific brand of condom
- number of condoms sold (of a social marketing brand)

Appendix B

Notes on Sampling

This manual provides general guidelines for the monitoring and evaluation of projects. It is not a textbook on sampling, however, it includes some basic principles related to sampling for quantitative studies and selection criteria for participants in qualitative studies.

Sampling for Quantitative Surveys

The need for sampling stems from the fact that it is not practical to study all the cases in a population (called “the universe” of cases) to measure trends within a given population. Although it is true that a national census tries to reach all the members of a population, the majority of studies look for a subgroup (or sample) of the universe. One exception occurs when there are very few cases or elements in the universe, and it is feasible to study all the cases (e.g., all the clinical facilities that offer HIV testing in a country in which few exist.)

When one uses service statistics to evaluate some aspect of a program, it is usually feasible to collect and analyze data for all cases, although it may not be necessary with extremely large

case loads. The need to take a sample generally applies to surveys of the target population.

The universe of a given study can consist of people, institutions, communities, or other statistical units. Some examples include:

- all the adults age 15-60 years old in a country;
- all the medical staff providing services to people living with HIV/AIDS in public sector facilities;
- all the clinics that offer HIV counseling and testing in the country;
- all the secondary school students from the public schools;
- all the commercial sex workers that work in established brothels in the capital city.

The objective of taking a sample is to select a subgroup that is representative of the entire universe, so that the results obtained from the sample provide an unbiased measure of the true level within the larger population (universe).

Two basic types of sampling exist:

1. probability sampling and
2. non-probability sampling.

Probability Sampling¹⁵

Probability sampling (when done correctly) yields results that are representative of the population under study in the rigorous

¹⁵ The description of sampling for quantitative analysis presented in this appendix draws heavily on Fisher and Foreit, *Designing HIV/AIDS Intervention Studies: An Operations Research Handbook*. Population Council. Washington, D.C., 2002.

sense of the word. Probability sampling is characterized by three criteria:¹⁶

- each element has some probability of being selected (although each element does not necessarily have the same probability);
- the probability of being selected is known; and
- the cases (units, respondents) are chosen at random.

Probability sampling reduces bias and yields results that can be generalized to the population under study.

The most common techniques of probability sampling are:

1. Simple Random Sampling

The evaluator assigns a unique number to each case or element. Using a table of random numbers or a lottery technique, he/she selects the desired number of cases or elements.

2. Systematic Sampling

In this case, the evaluator works with a list of possible cases or elements to determine the interval necessary to reach the desired sample size. For example, if 20,000 elements or cases exist in the universe and the investigator decides to take a sample of 500 cases, then he/she calculates the interval dividing the 20,000 by 500, which yields 40. The evaluator then takes every 40th case from the preestablished list, beginning at a randomly selected starting point.

¹⁶Adamchak et al. *A Guide to Monitoring and Evaluating Adolescent Reproductive Health Programs*. FOCUS on Young Adults Project. Washington, D.C., 2000.

3. Stratified Sampling

This type of sampling can be used with either of the two types described above, to ensure that different subgroups are represented in the final sample. For example, if one wanted to study the quality of services in health units in a region that has three hospitals, 30 health centers, and 60 health posts, it would be useful to stratify the three types of units to ensure that they are all represented in a final sample. In such a case, the evaluator establishes the number of units to be selected in each unit category, prepares a list of all the units by category, or stratum, and selects at random the predetermined number from each stratum. If the evaluator did not use stratified sampling, it would be possible to get a sample that would not include even one hospital.

An example of stratified sampling is lot quality assurance sampling (LQAS), which was developed for the manufacturing industry but is now commonly used in public health.

4. Multi-stage or Cluster Sampling¹⁷

This refers to selecting groups or clusters of elements, using simple random sampling or systematic sampling. For example, the Demographic and Health Survey (DHS) uses cluster sampling. The sampling plan generally is based on census areas, which constitute the clusters in this case. In the first stage, the evaluator or sampling expert randomly selects a set number of these census areas (clusters). In the second stage, within each selected area, he/she randomly selects a specific number of households. The interviewer then attempts to contact and interview each person within the selected household that

¹⁷For a simple but detailed explanation of this technique, see: Adamchak et al. *A Guide to Monitoring and Evaluating Adolescent Reproductive Health Programs*. FOCUS on Young Adults Project. Washington, D.C., 2000.

fits the selection criteria for the study, for example, women 15 to 44 years of age. This process is called “two-stage sampling” and is the most common type of cluster sampling.¹⁸

Cluster sampling is often less expensive than simple random sampling, because sampling is restricted to certain geographical areas. However, larger sample sizes are required to account for homogeneity within clusters.

Non-probability Sampling

Non-probability sampling does not yield samples that are truly representative of the universe. However, in practice, this approach is used under the following conditions:¹⁹

- The universe has less than 20 elements (e.g., clinical facilities that carry out HIV testing in a low prevalence country);
- It is difficult or impossible to construct a list of all the elements (e.g., the number of CSWs in the capital city).

Non-probabilistic sampling also is used at times for lack of human, financial, or technical resources. This sampling approach may yield some idea of the present trends in the target population. However, data collected with non-probability

¹⁸ More than two stages can exist. In the CDC (Centers for Disease Control and Prevention) surveys on reproductive health, if there is more than one eligible woman 15 to 44 years of age, the interviewer selects only one woman at random to be surveyed. This constitutes a third stage.

¹⁹ Adamchak et al. *A Guide to Monitoring and Evaluating Adolescent Reproductive Health Programs*. FOCUS on Young Adults Project. Washington, D.C., 2000.

sampling do not have the same value as those obtained through probabilistic sampling, since the findings cannot be generalized to the target population.

The four main types of non-probability sampling are:

1. Accidental or Convenience Sampling

The sample is drawn based on an opportunity that arises at the time of data collection (e.g., all the patients that come to a clinic on a given day).

2. Purposeful Sampling

The cases are selected for a specific reason (e.g., choosing one clinic in each area of the country — north, south, east, and west — to achieve geographic spread).

3. Quota Sampling

The characteristics of the sample are established beforehand (for example, a total number of cases is fixed at 200 women). The sample is established to reflect the distribution of the population: for example, 100 urban cases, 100 rural cases; from each group, 70 women with a primary education, 30 with a higher education. The quota sample cannot be considered “representative” (since not all the elements have a possibility of being selected). However, the resulting sample has characteristics somewhat similar to the population under study.

4. “Snowball” Sampling

Data are collected on a small group of people with special characteristics (e.g., men who have sex with men), and these respondents are asked to identify other people with similar characteristics to participate in the survey. The interviewers then contact the persons suggested, and upon completing these

interviews, they ask the respondents for other names, until they obtain an adequate number of cases or arrive at a point when no more new information about the topic is forthcoming.

Size of the Sample

The size of sample refers to the number of elements (e.g., people, health units) to be selected to ensure that the results of the evaluation are sufficiently precise. Frequently, the availability of resources and the anticipated use of the data (e.g., the statistical processes that the evaluator intends to use, once the data are collected) influence the decision.

From the start, it is important to distinguish between evaluations that measure the level of a phenomenon at a given moment and evaluations that measure results or change over time (e.g., before and after the project). Process evaluations frequently measure some phenomenon in a given moment, such as “satisfaction with project services.” In contrast, evaluations that track change measure one or more indicators over time; for example, the percentage of clients that demonstrate “correct use of a condom” before and after the intervention. The size of the sample needed to measure changes in indicators over time is greater than what is needed to measure the same indicator in a given point in time.

Box 51 presents the suggested sample size to measure an indicator at a given moment (Fitz-Gibbon and Morris, 1987). The authors explain that “N” in the table refers to the size of the total population (for example, the target population of the country), and the “s” indicates the suggested sample size. If one takes a sample of this size using random sampling, the results should be representative of the population. On the other hand, if one takes a smaller sample, the results are less reliable.

The more common and more complicated case involves the evaluation of change between pretest and posttest. To determine the size of the sample in this case, the evaluator follows these steps:

- Define the key indicator or indicators, based on the objectives of the project, that is expected to change (e.g., knowledge of modes of transmission, attitude towards people living with AIDS, use of condom in the last sexual encounter).
- Estimate (or “guesstimate”) the levels or values of these indicators in the target population prior to the intervention or project. (For example, results from previous research in this population might show that 50 percent know three methods of transmission, and barely 10% used a condom in their last sexual encounter).
- Define the magnitude of change considered adequate. For example, assuming 50% of the target population knows three methods of transmission, what does the evaluator consider to be an adequate amount of change: From 50 to 60? From 50 to 65? From 50 to 70?

**Box 51: Determining the Sample Size
for a Given Population²⁰**

N	s	N	s	N	s
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	228	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	106	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	133	1000	278	75000	382
210	136	1100	285	100000	384

²⁰ Original source: Krijcie, R.V. and D. W. Morgan, 1970. "Determining Sample Size for Research Activities." *Educational and Psychological Measurement*, 30, 607-610. Fitz-Gibbon and Morris (1987), page 163.

***Box 52: The Minimum Sample Necessary
to Measure Change Over Time***

P_1 (starting level of indicator)	P_2 (anticipated level to be achieved)	N (required sample level)
0.10	0.20	438
0.10	0.25	216
0.20	0.30	638
0.20	0.35	300
0.30	0.40	775
0.30	0.45	353
0.40	0.50	843
0.40	0.55	376
0.50	0.60	843
0.50	0.65	368
0.60	0.70	775
0.60	0.75	330
0.70	0.80	638
0.70	0.85	267
0.80	0.90	438
0.80	0.95	163

The sample sizes presented here assume a significance and power of 90 percent (levels commonly used with cluster sampling).

Box 52 indicates the sample size needed to measure changes over time, depending on the level of the indicator *before* the intervention and the level of change considered adequate.²¹ This table uses a significance level of 90 percent²² and a power of 90 percent.²³

The first column of Box 52, P_1 , indicates the estimated level (percentage/proportion) of the key indicator prior to the project²⁴ (e.g., the percentage that know three methods of transmission). The evaluator defines the magnitude of change needed to conclude that the project has achieved its objective (e.g., an increase from 50 to 65 percent). Observing the tenth line of numbers in the table, one can see that beginning with the anticipated value of .50 (50 percent know three methods of transmission) and anticipating an increase to .65 (65 percent will know three methods of transmission), it would be necessary to take a sample of 368 cases.

Given the importance of sampling, it is recommended that the project staff consult with a statistician or sampling expert before taking a sample for a survey. Box 52 represents only one of the possible options. For example, evaluators frequently get the significance level at 95 percent and the power at 80 percent,

²¹ That is, how large does the change in the indicator have to be for us to consider that the project has achieved its objective?

²² This implies a 90 percent certainty that the change has not occurred by chance.

²³ This implies a 90 percent certainty of detecting change if in reality it has occurred.

²⁴ Generally, one does not know with certainty what the value of the indicator will be before measuring it in the first survey (baseline). However, it is necessary to “guess” or to “anticipate” this number, based on previous surveys and other information sources.

which would change the number of cases necessary to measure change of a certain magnitude.

In this appendix we have presented some fundamental concepts of sampling, recognizing that this topic is complex and that these explanations are simplified. Nevertheless, this brief overview is intended to give readers the basic concepts of sampling, which should help to orient their discussions with local sampling experts.

Selection of Cases of Qualitative Studies

The rules of sampling have been developed to permit a certain level of precision in quantitative findings (e.g., percentages, averages) obtained from samples of the population.

However, in the case of qualitative research, the rules for selecting the number of cases are different, and they depend on the type of qualitative methodology used.

Observation (Direct and with Mystery Clients)

For this technique, the evaluator determines the “universe” of possible sites, situations, people, or cases to be observed as a first step in selecting the cases.

For example, let’s assume that we want to observe the quality of counseling services in province X of country Y.

- First, we have to determine how many facilities (e.g., hospitals, centers, laboratories) in the catchment area provide counseling services;
- Second, we determine how many staff in each facility give counseling.

-
-
- Third, we determine the number of counseling sessions to be observed.

One approach would be to select the number of observations to be carried out in proportion to the number of clients that attend each facility (for example, if a clinic carries out 75% of the counseling sessions, then 75% of the observations should come from this clinic) Alternatively, if one wants to compare the quality of HIV testing services in the five clinics of a country that provide the service, then one should carry out an equal number of observations in each clinic. Within each clinic, it is important to select the people to be observed, using the same type of criteria (proportional to the average number of clients that use the service or an equal number from each provider).

In this example the evaluator can apply the same rules described in the previous section to determine the sample size, given that direct observation generally consists of a process evaluation, and the measurement is taken at a given point in time.

In-depth Interviews

This technique is used to obtain detailed information on attitudes, opinions, and values of a small but carefully selected group of people.

This technique yields a large amount of information, given that the interviewer tries to capture all the information provided by the key informant. Because each transcription (each case) requires revision, analysis, and synthesis, the evaluator should be cautious in deciding the number to be carried out.

The number of respondents often ranges from 20 and 40. The selection depends on the objective of the research/evaluation. As Fisher *et al.* (1991) recommend, it can be useful to

intentionally select a variety of profiles to survey, including people that can give articulate responses to complex questions.

For example, to determine the community acceptance of educational sessions that discuss the topic of human sexuality with youth in public schools, one could interview teachers, parents, community leaders, religious leaders, and others that can give relevant opinions on the topic.

Another example: To learn more about the quality of care in a clinic serving people living with HIV/AIDS, one should conduct in-depth interviews with actual clinic users.

Focus Groups

One key component to focus groups is selecting appropriate respondents, based on characteristics defined by the evaluator, consistent with the objectives of the project. Often the evaluator will plan to conduct a series of focus groups to obtain the ideas and opinions of different subgroups. However the composition of any single group should be homogeneous with respect to characteristics relevant to the subject (e.g., age, sex, race, education, place of residence).

The number of groups to conduct depends on the available resources for evaluation (money and technical experience in conducting and analyzing results). However, as a general rule one should carry out at least two groups within each population subgroup to be covered. It is desirable, but not mandatory, that the participants not know one another, to avoid an established social hierarchy that could influence the dynamics of the group discussion.

To continue with the above example, suppose we want to evaluate the quality of care given to people living with HIV/AIDS in a specific clinic.

An important first question is the ethical aspects of the evaluation. The patients should be informed of their right not to participate if they do not want to, and that their refusal will not affect the service received.

The simplest approach could be for the evaluator to conduct a group discussion with patients who attend the clinic. However, it is important to consider whether participants would feel comfortable discussing certain topics with others. For example, if there are men and women in the same group, both groups might feel uncomfortable openly expressing their opinions. In this case, it would be advisable to conduct separate groups by sex. Gay men will express themselves more freely in a group consisting only of gay men. Sex workers will generally feel more comfortable discussing topics of interest with other sex workers. Within these groups, it may be useful to define a limited age range, to give participants a further level of comfort in openly discussing issues.

In the case under discussion, the evaluator might decide to conduct separate groups for:

- men who have sex with men (20-29, 30-39 years)
- commercial sex workers (20-29, 30-39 years)
- men who self-identify as gay (20-29, 30-39 years)
- women (e.g., homemakers, working women) (20-29, 30-39 years)

In deciding the total number of groups to conduct, the evaluator should take into consideration whether the analysis will be done manually or with a computer software package. When computer analysis is not an option, we recommend limiting the number of groups to eight or less.

In this appendix we have approached the basic questions related to sampling from the point of view of organizations that do not have extensive experience in research and evaluation and/or that have limited resources. Consulting with a statistical expert is always useful when making decisions about sampling. In quantitative research, a poor sample will not yield valid results, since the findings will not be representative of the population under study. In qualitative research, careful selection of participants makes the findings more useful for programmatic purposes.

Appendix C

Additional Bibliographic References on Topics Covered in This Manual

Chapter 2: Concepts and Definitions

Behets, F. et al. "Estimating the Cost and Effectiveness of Different STI Management Strategies for Sex Workers in Madagascar." Horizons Research Summaries, 2002. (Available on the Internet at: <http://www.popcouncil.org/horizons/ressum/swmad/swmad.html>)

Creese, A. and D. Parker. Cost Analysis in Primary Health Care: A Training Manual for Programme Managers. England: World Health Organization, 1994. (Available on the Internet at: <http://whqlibdoc.who.int/publications/9241544708.pdf>)

Fisher, A.A, J.E. Laing, J.E. Stoeckel, J.W. Townsend. Handbook for Family Planning Operations Research Design. New York: The Population Council, 1991.

Fitzgibbon, C. Taylor and L.L. Morris. How to Design a Program Evaluation. Newbury Park, CA: Sage Publications, 1987.

Franco, X., L. Miller, P. Richardson, J. Reynolds, and N. Kak. Monitoring and Evaluating Programmes. Module 5 User's Guide. Washington, DC and Geneva, Switzerland: The Aga Khan Foundation, 1993.

Green, L. and F.M. Lewis (eds), Measurement and Evaluation in Health Education and Health Promotion. Palo Alto, CA: Mayfield Publishing Company, 1986.

Hogle, J. and M. Sweat. FHI/AIDSCAP Evaluation Tools: Qualitative Methods for Evaluation Research in HIV/AIDS Prevention Programming. Research Triangle, NC: Family Health International, 1996.

Krause, D. Effective Program Evaluation, An Introduction. Chicago, Ill: Nelson-Hall Publishers, 1996.

Rossi, P., H. Freeman and M. Lipsey. Evaluation: A Systematic Approach. Thousand Oaks: Sage Publications, 1999.

Chapter 3: Designing an Evaluation

Aubel, J. Participatory Program Evaluation Manual: Involving Program Stakeholders in the Evaluation Process. Calverton, MD: ORC Macro, 2000. (Available on the Internet at: <http://www.childsurvival.com/documents/CSTS/ParticipatoryEvaluation.cfm>)

Bertrand, J.T., R.J. Magnani, and N. Rutenberg. Evaluating Family Planning Programs with Adaptations for Reproductive Health. Chapel Hill, NC: Carolina Population Center. The EVALUATION Project. 1996. (Available in Spanish on the Internet: <http://populi.eest.ucr.ac.cr/manuscritos/manual/index.htm>)

Fisher, A. and J.R. Foreit. Designing HIV/AIDS Intervention Studies: An Operations Research Handbook. New York: Population Council, 2002. (Available on the Internet at: <http://www.popcouncil.org/horizons/orhivaidshndbk.html>)

Mertens, T., M. Carael, P. Sato, J. Cleland, H. Ward and G.D. Smith. "Prevention indicators for evaluating the progress of national AIDS programmes." AIDS 8:1359-69, 1994.

Chapter 4: Methodologies for Collecting Data

Bernard, H.R. *Research Methods in Anthropology*, 2nd edition, Walnut Creek, CA: Altamira Press/Sage Publications, 1994.

Debus, Mary. *The Handbook for Excellence in Focus Group Research*. Washington, DC: Academy for Educational Development, 1998. (Available in Spanish under the title: *Manual para Excelencia en la Investigación Mediante Grupos Focales*)

Ferencic, N. *Guide for Carrying Out In-Depth Interviews about Health in Developing Countries*. Philadelphia, PA: University of Pennsylvania, 1988.

Nelson, K., L. MacLaren, R. Magnani. (2000). "Assessing and Planning Youth-Friendly Reproductive Health Services." *Focus on Young Adults Tool Series No. 2. Focus on Young Adults*, Washington, DC, 2000.

Patton, M.Q. *Qualitative Evaluation and Research Methods*, 2nd edition, Newbury CA: Sage Publications, 1990.

Van Sant, Jerry. "Qualitative analysis in development evaluations." *Evaluation Review* 13(3): 257-72, 1989.

Chapter 5: HIV/AIDS Illustrative Projects

Academia para el Desarrollo Educativo (AED), La Cadena de Cambios Cuatro Pasos para Fortalecer su Programa, Ciudad de Guatemala, Guatemala: AED, 1997.

The Evaluation Project, Indicators for Reproductive Health Program Evaluation. Final Report of the Subcommittee on Adolescent Reproductive Health Services, Chapel Hill, Carolina Population Center, 1995.

UNAIDS, Summary Booklet of Best Practices. Geneva, Switzerland: UNAIDS, 1999.

UNAIDS, Sexual behavioral change for HIV: Where have theories taken us? Geneva: UNAIDS, 1999.

UNAIDS, National AIDS Programmes: A Guide to Monitoring and Evaluation Geneva UNAIDS, 2000.

Chapter 6: Data Processing and Analysis

Fitz-Gibbon, C.T. and L.L. Morris. How to Analyze Data. Newbury Park, CA: Sage Publications, Ltd., 1987.

García-Nuñez, J. Evaluación de Programas de Planificación Familiar: Guía Detallada para Administradores y Evaluadores. Santa Fé de Bogotá, Colombia. Asociación Pro-Bienestar de la Familia Colombiana (PROFAMILIA), 1992

Zeller, R.A. and E.G. Carmines. Statistical Analysis of Social Data. Chicago, Ill: Rand McNally College Publishing Company, 1978.

