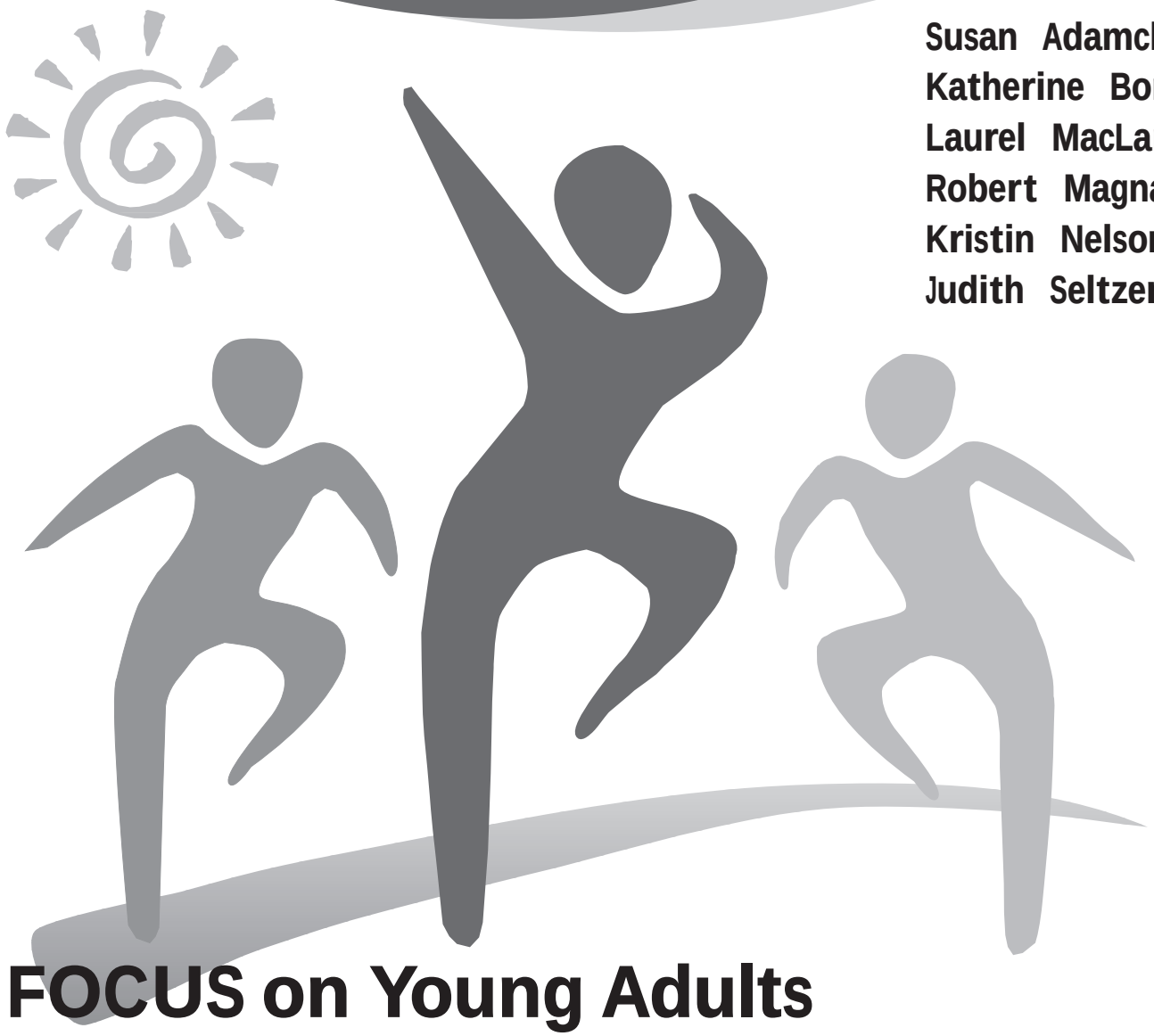
A GUIDE TO

Monitoring and Evaluating Adolescent Reproductive Health Programs

Susan Adamchak
Katherine Bond
Laurel MacLaren
Robert Magnani
Kristin Nelson
Judith Seltzer



FOCUS on Young Adults

Tool Series 5, June 2000

The FOCUS on Young Adults program promotes the well-being and reproductive health of young people. FOCUS is a program of Pathfinder International in partnership with The Futures Group International and Tulane University School of Public Health and Tropical Medicine. FOCUS is funded by USAID, Cooperative Agreement # CCP-A-00-96-90002-00. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Agency for International Development.

Please send suggestions or comments to:

FOCUS on Young Adults
Attn: Communications Advisor
1201 Connecticut Avenue, NW, Suite 501
Washington, DC 20036, USA

Tel: 202-835-0818
Fax: 202-835-0282
Email: focus@pathfind.org

The authors are indebted to the many people who contributed to the development and review of this Guide.

We wish to acknowledge the dedicated efforts made by several graduate research assistants working with FOCUS on Young Adults at the Tulane University School of Public Health and Tropical Medicine, Department of International Health and Development. Stephanie Mullen began the detailed project of compiling program indicators. Gwendolyn Morgan prepared the appendices listing recommended references and Internet Web sites, and provided formulae for the Indicator Tables. Emily Zielinski assisted with the Indicator Tables and appendices.

Our FOCUS colleagues, Sharon Epstein, Lindsay Stewart, Barbara Seligman and Lisa Weiss, read early versions of this Guide and offered helpful suggestions. Their comments reminded us to keep in the forefront of our efforts the many program staff we hope will find this volume useful.

The authors would like to express their appreciation to FOCUS staff member Christine Stevens for her critical review and recommendations for reorganizing several chapters of the Guide. We would also like to recognize Laura Sedlock, whose accomplished editing did much to clarify concepts and blend the voices of the authors.

Ideas and concepts that shaped the development of this Guide were discussed at a FOCUS Research and Evaluation working group meeting in April 1998. Those who participated in the discussion included Lisanne Brown (Tulane University), Nicola Bull (UNICEF), James Chui (UNFPA), Richard Colombia (Pathfinder International), Bruce Dick (UNICEF), Jane Ferguson (World Health Organization), Alix Grubel (John Snow International), Paula Hollerbach (Academy for Educational Development), Marge Horn (USAID), Merita Irby (International Youth Foundation), Lily Kak (CEDPA), Rebecka Lundgren (Georgetown Institute for Reproductive Health), Matilde Maddaleno (Pan American Health Organization), Leo Morris (Centers for Disease Control), Lisa Mueller (John Snow International), Ian Tweedie (Johns Hopkins University Center for Communications Programs), Stephanie Mullen (Tulane University), Phyllis Scattergood (Education Development Center, Inc.), Annetta Seecharan (International Youth Foundation), Linda Sussman (USAID), Katherine Weaver (Pan American Health Organization), Ellen Weiss (Population Council/Horizons) and Anne Wilson (PATH).

Those who provided critical comments and feedback during the field review of this Guide include Jane Bertrand (Tulane University), Carlos Brambila (Population Council, Mexico), Eunyong Chung (USAID), Charlotte Colvin (The Futures Group International), Shanti Conly (USAID), Barbara deZalduondo (USAID), Joyce Djaelani (PATH Indonesia), Maricela Dura (Fundación Mexicana para la Planeaciùn Familiar), Natalia Espinoza (CEMOPLAF Ecuador), Julie Forder (CARE Cambodia), Phyllis Gestrin (USAID), Evam Kofi Glover (Planned Parenthood Association of Ghana), Y.P. Gupta (CARE India), Lisa Howard-Grabman (Save the Children), Douglas Kirby (ETR Associates), Rekha Masilamani (Pathfinder International, India), Ruth Maria Medina (Population Council, Honduras), Dominique Meekers (Population Services International), Irene Moyo (JSI/SEATS), Nancy Murray (FOCUS on Young Adults), Mary Myaya (CARE Lesotho), Sonia Odria (Pathfinder International, Peru), Oladimeji Oladepo (Department of Preventive and Social Medicine, Nigeria), Anne Palmer (PATH

Philippines), Susan Pick de Weiss (Instituto Mexicano de Investigación de Familia y Población), Gabriela Rivera (Pathfinder International, Mexico), William Sambisa (PACT Zimbabwe), Jessie Schutt-Aine (International Planned Parenthood Federation), Alfonso Sucrez (Fundación Mexicana para la Planeación Familiar), Oswaldo Tanako (Pan American Health Organization), John Townsend (Population Council/Frontiers), Laelani L.M. Utama (Pathfinder International, Indonesia), Pilar Vigal (CEBRE, Chile), Amy Weissman (Save the Children), Anne Wilson (PATH) and Kate Winskell (Global Dialogues).

Presentations and participant discussion at the YARH Measurement Meeting sponsored by the Centers for Disease Control (CDC) Division of Reproductive Health and FOCUS on Young Adults in September 1999 helped shape the discussion of data collection. In particular, presentations by Gary Lewis (Johns Hopkins University Center for Communications Programs), Paul Stupp (CDC Division of Reproductive Health) and Cynthia Waszak (Family Health International) were helpful in finalizing this Guide.

Health and Human Development Programs staff of the Education Development Center, Inc. (EDC), managed the review process under the able direction of Phyllis Scattergood and Carmen Aldinger.

EDC's Editing and Design Services, led by Jennifer Roscoe, was responsible for the production of this Guide, including design and coordination by Cathy Lee and revisions and copyediting by the editorial staff. Their creative input is very much appreciated.

Acronyms and Abbreviations

ARH	adolescent reproductive health
BCC	behavior change communication
CEA	census enumeration area
DHS	Demographic and Health Survey
FLE	family life education
IEC	information, education and communication
M&E	monitoring and evaluation
MIS	management information system
MOS	measure of size
NGO	nongovernmental organization
PPS	probability-proportional-to-size
RH	reproductive health
RTI	reproductive tract infection
STD	sexually transmitted disease
STI	sexually transmitted infection
USAID	United States Agency for International Development

About the Authors

Susan E. Adamchak is president of Planning & Evaluation Resources, Inc. Her areas of expertise include population and health policy development, program assessment and evaluation of reproductive health and public health programs. She holds a PhD in Sociology from Brown University.

Katherine Bond is Research Assistant Professor at the Tulane University School of Public Health and Tropical Medicine, and Research and Evaluation Advisor at FOCUS on Young Adults. She has managed HIV/AIDS programs for youth in the United States and Thailand, and has trained governmental and nongovernmental organizations in Asia and Africa on the use of social research methods for program design and evaluation. She has a doctorate in international health from Johns Hopkins University.

Laurel MacLaren was the Communications Coordinator at FOCUS on Young Adults. She founded and managed an adolescent sexual health program with the Indonesia Planned Parenthood Association in Yogyakarta and has provided technical assistance on adolescent reproductive health program design, monitoring and evaluation in South and Southeast Asia. She has a master's degree in public policy from Harvard University.

Robert J. Magnani, PhD, is currently an Associate Professor in the Department of International Health and Development of the Tulane School of Public Health and Tropical Medicine. He has worked in the international population and health fields in the areas of data collection systems and methodology, program/project monitoring and evaluation, and information systems support to program management and policy decisionmaking. He has worked in 27 developing countries in all regions of the world, with specialization in East/Southeast Asia and Latin America.

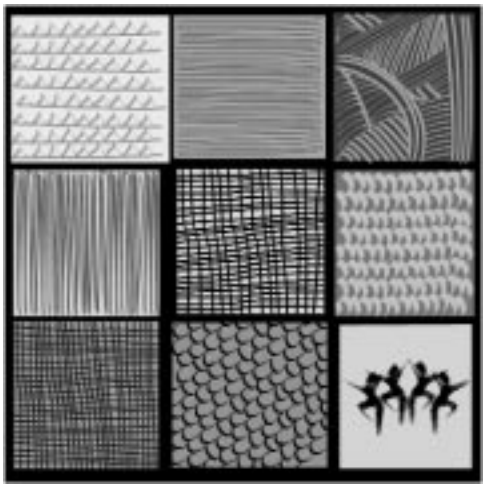
Kristin Nelson is a DrPH candidate at Tulane University and has a master's degree in medical anthropology from Case Western University. She has conducted extensive review of qualitative and quantitative instruments and youth programs for FOCUS on Young Adults. She lived and worked in Tanzania for two years and has experience working in AIDS education for youth in Ethiopia.

Judith R. Seltzer is an independent consultant and population specialist with an emphasis on population policy, family planning and reproductive health, and design and evaluation of international population assistance programs. She has a PhD from Johns Hopkins University.

PHOTO: JHU/CCP

TABLE OF CONTENTS

PART II: INSTRUMENTS



INTRODUCTION

Around the world, young people are growing up in an environment of dynamic change. For some, this complexity provides opportunity and choice; for others, it means a struggle for survival. Many young people have stamina and energy, curiosity, a sense of adventure and invulnerability. They are resourceful and resilient even under the most difficult conditions.

The period of adolescence is, however, a life phase in which young people are particularly vulnerable to health risks, especially those related to sexuality and reproduction: HIV/AIDS, unwanted pregnancy, unsafe abortion, too-early marriage and childbearing, sexually transmitted infections and poor nutrition.

How adolescence is experienced and affects reproductive health has largely to do with the timing and sequence of sexual initiation, marriage and childbirth; the degree to which the timing and sequence of these events are socially sanctioned or forbidden; and the number and availability of options regarding education, job training and employment. There is a great deal of variation worldwide, and even within countries, in the social and cultural values that shape these events. Close relationships between youth and their parents and extended family are particularly important in influencing youth development. Access to preventive and curative services, including contraception and treatment for sexually transmitted infections, are also important in ensuring the reproductive health of youth.

Youth development programs designed to help young people reduce their reproductive health risks reflect that variation. Many of these programs regard young people as a critical resource for the future, and use creative strategies to tackle their complex problems. But many programs face limited funding, community resistance, nonsupportive laws and policies or lack of experience. By knowing more about what works in youth programs and services, we can build strong programs that accomplish what they intend.

Reproductive health refers to the health and well-being of women and men in terms of sexuality, pregnancy, birth and their related conditions, diseases and illnesses. Many programs reaching youth are trying to achieve reproductive health goals that relate to critical sexual and reproductive health outcomes, such as:

- fertility: the number of pregnancies a woman has in her lifetime
- abortion: as it relates to fertility and to health complications for women who have unsafe or clandestine abortions

- illness: caused by sexually transmitted infections, reproductive tract infections, HIV and/or nutritional status
- mortality: primarily related to pregnancy and childbearing, including infant and maternal mortality, and also including AIDS-related deaths
- nutritional status: which impacts both women's health and that of their infants

Note

What do we mean by "youth"?
Programs reaching young people use different terminology to refer to youth. "Adolescents" is often used to refer to young people ages 10–19, "young adults" generally refers to those ages 15–24 and "youth" may refer to all young people ages 10–24. This guide encompasses each term and uses the phrase "adolescent reproductive health" (ARH) to cover each type of program.

Why Monitor and Evaluate Youth Programs?

Monitoring and evaluation shows if and how youth programs are working.

Monitoring and evaluation (M&E) can tell us if and how program activities are working. Program managers and donors want to be able to demonstrate results, understand how their programs are working and assess how the programs interact with other events and forces in their communities.

M&E can be used to strengthen programs.

Program managers and staff can assess the quality of activities and/or services and the extent to which the program is reaching its intended audience. With adequate data, you can compare sites, set priorities for strategic planning, assess training and supervisory needs and obtain feedback from the target audience or program participants. You can prioritize resource allocation, improve information for fund-raising, provide information to educate and motivate staff, provide information for advocacy and argue for the effectiveness of your program approach.

M&E results can help institutionalize programs.

M&E results can help stakeholders and the community understand what the program is doing, how well it is meeting its objectives and whether there are critical needs inhibiting your progress. M&E results can be used to educate your board of directors, current and prospective funding agencies, local government officials and key community members—such as local leaders, youth and parents—who can help ensure social, financial and political support for youth programs. Sharing results can help your program establish or strengthen the network of individuals and organizations with similar goals of working with young people. It can also give public recognition and thanks to stakeholders and volunteers who have worked to make the program a success, and may attract new volunteers.

M&E shapes the decisions of funding agencies and policymakers.

Funding agencies and policymakers are interested in monitoring and evaluation results for a variety of reasons. They need to make strategic choices about how to spend resources and to prove that the expenditure produces quality results. M&E results also help with decisions about identifying and supporting the replication or expansion of particular program strategies. M&E findings often reveal unmet needs or barriers to program success and can be used to lobby for policy or legislative changes. M&E results can raise awareness of youth programs among the general public and help build positive perceptions about young people and youth programs.

M&E results contribute to the global understanding of "what works."

The dissemination of M&E results—both those that show how your program is working and those that find that some strategies are not having the intended impact—contributes to our global understanding of what works and what doesn't in improving young people's reproductive health. This advances the field by building a body of lessons learned and best practices that can strengthen ARH programs around the world.

M&E mobilizes communities to support young people.

Monitoring and evaluation results enable communities and youth to inform local leaders about youth needs and to advocate for funding. Results point to ways in which we can develop new and better systems of support for young people and identify additional community resources. They can increase the community's understanding of the potential and actual benefits of the program and its accomplishments, develop a sense of ownership through participation, improve coordination and mobilize support for youth and the array of programs that foster their health and development.

Who Should Use This Guide?

This Guide is designed for program managers who monitor and evaluate adolescent reproductive health programs. Some examples of the people who might find this guide useful include the following:

- **Community-level program managers:** A manager of a community youth center's peer education program can use this Guide to set up a system to monitor implementation of program activities.
- **District-level program directors:** A director of a school-based family life education (FLE) program can use this Guide to track progress in the program's implementation.
- **Municipal-level health managers:** A manager of a clinic's pregnancy and sexually transmitted infection (STI) reduction program can use this Guide to set up an evaluation that will track changes in the incidence of pregnancy and STIs among youth in the entire municipality.
- **State- or provincial-level health officials and managers of nongovernmental organizations (NGOs):** An official at the state level in a health system can use this Guide to compile data across districts, municipalities or other geographic areas or population groups to develop a picture of the current status of youth health, as well as changes over time.
- **Managers or technical staff of private voluntary or donor agencies:** A manager of a private voluntary agency can use this Guide to advise other organizations on how to improve their programs and how to set up a monitoring and evaluation system for youth programs.

Note

Seeking outside help

Monitoring and evaluation is an essential aspect of youth reproductive health program development. However, many programs do not have the expertise to carry out some aspects of program evaluation, especially when evaluating large, complex programs. After reading this Guide, you may choose to seek technical assistance from local universities and research institutes who have the expertise to help you design and conduct an effective and efficient evaluation.

Origins of This Guide

This Guide draws on the expertise and experience of professionals in a variety of disciplines.

The family planning field has laid an important foundation for considering how to develop service delivery systems for adults and how to measure inputs, quality, access and program results. This Guide draws heavily on the contributions of USAID's The EVALUATION Project, which approaches evaluation with a focus on a program's systems and delivery and an extensive menu of reproductive health outcome indicators.

M&E results can help stakeholders and the community understand what the program is doing, how well it is meeting its objectives, and whether there are critical needs inhibiting your progress.

This Guide also draws lessons from the field of HIV/AIDS prevention, with its open view of sexuality and sexual behavior and its understanding of the value of social and behavioral change theory in designing effective programs for young people.

The youth development field, which has identified a range of developmental needs and assets, urges us to measure social influences beyond individual knowledge, attitudes and practices, such as building healthy relationships and supportive communities and fostering skills development.

The FOCUS on Young Adults program's own contributions in reviewing youth program experiences in developing country settings are incorporated in this Guide. Those reviews have contributed to our presentation of "key elements" of program design and possible criteria for establishing measures of program quality and access.

What Are Monitoring and Evaluation?

Monitoring and process evaluation measure how a program is working.

Monitoring is the routine tracking of a program's activities by measuring on a regular, ongoing basis whether planned activities are being carried out. Results reveal whether program activities are being implemented according to plan, and assess the extent to which a program's services are being used.

Process evaluation should be done along with monitoring. Process evaluations collect information that measures how well program activities are performed. This information is usually collected on a routine basis, such as through staff reports, but it may also be collected periodically in a larger-scale process evaluation effort that may include use of focus groups or other qualitative methods. Process evaluation is used to measure the quality of program implementation and to assess coverage; it may also measure the extent to which a program's services are being used by the intended target population.

Outcome and impact evaluation measure a program's result and effects.

Outcome and impact evaluation measure the extent to which program outcomes are achieved, and assess the impact of the program in the target population by measuring changes in knowledge, attitudes, behaviors, skills, community norms, utilization of health services and/or health status. *Outcome evaluation* determines whether outcomes that the program is trying to influence are changing in the target population. *Impact evaluation* determines how much of the observed change in outcomes is due to the program's efforts.[1]

This Guide has two parts, which are described below.

PART I: THE HOW-TO'S OF MONITORING AND EVALUATION

Chapter 1: Concerns About Monitoring and Evaluating ARH Programs

- ➤ Reviews challenges to and offers tips on measuring the effectiveness of youth programs
- ➤ Discusses how to be sure that your results are attributable to the program effort
- ➤ Previews ways this Guide can provide information and offer support

Chapter 2: A Framework for ARH Program Monitoring and Evaluation

- ➤ Considers the multiple factors that shape adolescence
- ➤ Introduces three major strategies used to improve youth reproductive health
- ➤ Discusses the Logic Model, an approach to designing an effective strategy

This Guide is designed for program managers who monitor and evaluate adolescent reproductive health programs.

Chapter 3: Developing an ARH Monitoring and Evaluation Plan

- ➤ Defines program goals, outcomes and objectives
- ➤ Helps you define the scope of your monitoring and evaluation effort
- ➤ Offers guidance on how to plan and conduct a monitoring and evaluation effoct, using the rest of this Guide

What Can You Determine Using Monitoring and Evaluation?

Monitoring & Process Evaluation	Outcome & Impact Evaluation
➤ Whether program is being implemented according to plan ➤ Quality of program ➤ Coverage of program	➤ Changes in outcomes, such as: – changes in behavior – changes in knowledge and attitudes – changes in interactions with parents – changes in community norms ➤ Whether outcomes are due to program efforts or other factors

Chapter 4: Indicators

- ➤ Defines and explains indicators
- ➤ Provides examples of how to select and modify indicators to match your program objectives and activities

[1] Outcome evaluations often measure short-term changes, such as changes in knowledge, attitudes and behaviors. Impact evaluations are often conducted over a longer period and are able to identify changes in sexual and reproductive health outcomes in the target population, such as rates of STIs.

Chapter 5: Evaluation Designs to Assess Program Impact

- ➤ Offers guidance on and considerations around the need for impact evaluation
- ➤ Reviews study designs you can use to carry out an impact evaluation
- ➤ Outlines the technical requirements and resources needed for each type of evaluation
- ➤ Presents options for initiating evaluations after a program is underway

The information you collect through monitoring and process evaluation will also help you build the case that the changes were a result of your program, even if an impact evaluation is not feasible.

Chapter 6: Sampling

- ➤ Describes types of sampling methods and ways to determine which one is appropriate for your program
- ➤ Focuses on one commonly used sampling method: cluster sampling
- ➤ Reviews how to determine and calculate the sample size you need for your program

Chapter 7: The M&E Workplan and Data Collection

- ➤ Reviews data collection steps
- ➤ Addresses ethical concerns
- ➤ Presents options for data collection methods
- ➤ Discusses tasks involved in developing an M&E workplan

Chapter 8: Analyzing M&E Data

- ➤ Details how to process both quantitative and qualitative data
- ➤ Reviews mechanics of data analysis
- ➤ Discusses how to analyze and interpret data to draw conclusions about program design, functioning, outcomes and impact

Chapter 9: Using and Disseminating M&E Results

- ➤ Reviews reasons to use and disseminate M&E results
- ➤ Describes how to use M&E results to improve your program's interventions
- ➤ Offers tips on how to disseminate results to priority target audiences
- ➤ Presents different formats for dissemination of results

Chapter 10: Tables of ARH Indicators

- ➤ Presents four tables of ARH indicators
- ➤ Features indicators for each phase of a program (program design, program systems development and functioning, program implementation and program intervention outcomes)
- ➤ Describes how to use the Indicator Tables

Glossary

Bibliography

Appendices

- ➤ Sampling schemes for core data collection strategies
- ➤ Calculating sample size requirements
- ➤ Reference shelf of useful books
- ➤ Relevant Internet sites

PART II:
INSTRUMENTS AND QUESTIONNAIRES

- ➤ Offers guidance on adapting instruments for your M&E effort
- ➤ Provides sample data instruments
- ➤ Gives tips for collecting data through a variety of methods

PHOTO: Harvey Nelson

CONCERNS ABOUT MONITORING AND EVALUATING ARH PROGRAMS

CHAPTER AT A GLANCE

- ➤ Reviews challenges to and offers tips on measuring the effectiveness of youth programs
- ➤ Discusses how to be sure that your results are attributable to the program effort
- ➤ Previews ways this Guide can provide information and offer support

Fifteen Challenges in Monitoring and Evaluating Youth Programs

1. Some MIS are not set up to track the special characteristics of youth programs.

Some MIS are part of a larger program or service delivery intervention. For example, a family planning program that has a youth component may be set up to track the distribution of contraceptives; it may not be set up to track services that are more likely to be utilized by youth, such as counseling or distribution of information, education and communication (IEC) materials.

Adapting your MIS to monitor an ARH program may require only minor modification, such as adding the specification of age in program utilization reporting. However, for larger-scale programs that reach groups other than youth, adding even one new component to the system may be difficult to institutionalize.

2. Tracking services does not guarantee that you will know how many youth you are reaching.

All programs need to determine how they will count the youth they are reaching and how knowing the number of youth reached will improve performance. Many programs count services, such as the number of meetings held or the number of condoms distributed. However, if all you know is that you distributed 1,000 condoms, you will not know whether 100 youth received 10 condoms each or 500 youth received 2 condoms each. Your information tracking system should try to collect key characteristics of program participants to help assess whether the program is reaching the number and type of youth it was designed to reach.

Collecting information about target population characteristics will also help you understand how your program participants change over time. For example, in the beginning, your program may target older youth, but as word spreads about the services available, your program may find

15 Challenges in Monitoring and Evaluating Youth Programs

1. Some MIS are not set up to track the special characteristics of youth programs.
2. Tracking services does not guarantee that you will know how many youth you are reaching.
3. You may be unsure whether general standards or implementation strategies are applicable in the country you work in.
4. Little is known about whether standards for adult programs are appropriate for youth.
5. The elements of successful youth programs have not been well-documented or disseminated.
6. Programs may have trouble developing systems that understand and respond to the needs of youth.
7. Measuring the quality of a program requires understanding complex meanings and addressing sensitive issues.
8. Measuring a program's access and coverage can be complex.
9. Assessing individual reactions to a program can be difficult.
10. Measuring influences on behaviors that didn't occur is difficult.
11. Measuring behaviors at a variety of developmental levels can be problematic.
12. Showing the link between health outcomes and youth development can be complex.
13. Some changes may not be measurable for a long time, and others may be hard to measure at all.
14. Attributing changes in outcomes to a particular program's strategy and activities is difficult.
15. Some types of evaluation are costly and may require funds beyond a youth program's resources.

itself working with younger adolescents and need to adjust its approach accordingly.

3. You may be unsure whether general standards or implementation strategies are applicable in the country you work in.

Quality refers to the appropriateness of a specific set of professional activities in relation to the objectives they are intended to serve.[1] Standards of quality for the design of health education programs have been drawn from a variety of youth programs demonstrated to be effective in changing specific behaviors[2] and include factors such as:

- ➤ a minimum of 14 hours of instruction,
- ➤ small groups and an interactive environment, and
- ➤ models of and practice in communication, negotiation and other skills.

However, we do not know the extent to which these standards apply in a more diverse set of developing country settings.

The recommendations in this Guide, such as the Logic Model described in Chapter 2, are designed to help you implement your program strategy, based on

[1] Green and Lewis, 1986.
[2] Kirby et al., 1997.

assumptions about the social and behavioral factors that influence the health outcomes you hope to produce. The theories these recommendations draw on are well-developed and have been through a rigorous process to test how well their measurements capture the processes of change they propose. Yet most of these theories have not been tested in developing country settings and need to be adapted to the particular needs of youth in each locale. Since program activities drive the design of any evaluation effort, our lack of understanding about how these theories apply in different contexts can also affect our ability to undertake solid outcome and impact evaluations.

4. Little is known about whether standards for adult programs are appropriate for youth.

After years of developing contraceptive service delivery systems for adults, there are now more or less accepted standards of quality. For example, there is wide consensus that the delivery of quality clinical contraceptive services entails:

- ➤ technical competence of service providers,
- ➤ respectful treatment of clients,
- ➤ effective communication with clients,
- ➤ choice of methods,
- ➤ mechanisms to encourage continuity, and
- ➤ cultural appropriateness and acceptability of services.[3]

However, we still do not know how comprehensive these standards are for younger age groups. Some of these quality standards are listed in the Indicator Tables as examples of criteria to include in indicators of quality, especially at the design stage.

5. The elements of successful youth programs have not been well-documented or disseminated.

Youth program staff in developing countries often must rely on intuition and experience to design their programs when they don't have access to documented research. However, much is known about the standards that produce effective programs. For example, the FOCUS on Young Adults program has identified the following "key elements":[4]

- ➤ baseline assessment conducted to identify issues, needs and target audiences;
- ➤ existence of a clearly defined mission statement that contributes to the achievement of program goals; and
- ➤ local stakeholders involved in program planning.

6. Programs may have trouble developing systems that understand and respond to the needs of youth.

Program systems and their functioning will influence factors such as staff performance, service delivery and program utilization. Program systems must be set up to respond to the special needs of young people. For example, the staff recruitment and training system must ensure that staff hold the characteristics and skills to which youth respond well. A program system will help identify whether program materials are being updated often enough to respond to the changing language and trends of youth culture. A training system must ensure that the necessary components of youth programming are included in the curricula.

[3] Bruce, 1990.

[4] Birdthistle and Vince-Whitman, 1997; Israel and Nagano, 1997; Senderowitz, 1997a; and Senderowitz, 1997b. Note that these key elements reflect the experiences of programs that are con cerned more with reproductive health outcomes than with youth development outcomes.

7. Measuring the quality of a program requires understanding complex meanings and addressing sensitive issues.

To determine program quality, you will probably have to elicit subjective interpretations, perspectives and meanings from young people and others in the community. These are each complex because they are based on:

- ➤ cultural beliefs and values,
- ➤ personal interactions within a community,
- ➤ interactions between the young people and the program's staff, and
- ➤ opinions and views of people carrying out the program.

Programs that are concerned with youth empowerment, community mobilization, changing social norms and influencing youth culture will need to explore the meanings of such issues as feelings of self-worth, the value of community connectedness and the interpretation of culture. These reflections may be difficult to elicit and harder still to quantify. For example, you may be able to count the number of community members at a meeting, but have more difficulty assessing their substantive contribution to the meeting, increased concern as a result of the meeting or proposed strategy for social change.

Your information tracking system should try to collect key characteristics of program participants to assess whether the program is reaching the number and type of youth it was designed to reach.

Substantive changes in meanings and perceptions are extremely important for youth programs and should not be minimized. They play an important role in the quality of a youth program. To capture these nuances, we need to first employ qualitative approaches to data collection. Once we understand the relevant meanings, values and beliefs we can then collect data about changes in the number of participants who share those meanings, values and beliefs, i.e., a quantitative approach.

There are numerous obstacles to measuring the outcomes of youth development and reproductive health programs, which helps explain why we have such a limited body of evidence as to "what works." First, many of the intended outcomes are regarded as personal and private. In some societies, talking about sexual behavior and personal relationships may be socially prohibited. Second, evaluators may face parental and community resistance to asking young people questions. Community leaders or other key stakeholders may believe that the young people in their communities do not engage in risky behaviors, and therefore there is no need to ask questions. They may also find it socially or politically dangerous to uncover the truth about young people's sexual behavior, and make an attempt to block data collection. However, there are many examples of programs that asked sensitive questions and found young people who were eager to discuss issues of sexuality and reproductive health—viewing the discussions as an opportunity for

learning and for sharing their own concerns and needs.

8. Measuring a program's access and coverage can be complex.

Access to reproductive health programs concerns the extent to which youth can obtain appropriate reproductive health services at a level of effort and cost that is both acceptable to and within the means of a large majority of youth in a given population.[5] We can define access in a variety of ways:

- **Geographic/Physical:** Convenient hours and location, wide range of necessary services
- **Economic:** Affordable fees
- **Psycho-social:** Perception of privacy; perception that both males and females, married and unmarried youth, are welcome; feeling of safety and confidentiality; perception that providers are interested in, informed about and responsive to youth needs
- **Administrative:** Specially trained staff with respect for young people, adequate time for interactions, youth involvement in design and continuing feedback, short waiting times

Coverage refers to the extent to which your program's services—such as educational or clinical services—are being used by your intended target population. Coverage can be measured by:

- determining the proportion of the target population you are reaching, or
- determining the characteristics of the population you are reaching.

Some aspects of accessibility and coverage can be measured by the absence or presence of something and may be relatively straightforward. For example, finding out whether your program has convenient hours and affordable fees may be easily determined with a short survey of your target population. However, measuring more subjective issues that involve judgments—such as whether staff have respect for young people—can be more difficult because many youth may be reluctant to give their true opinions about program staff for fear of negative consequences, such as having services withheld.

Similarly, determining some characteristics of youth may be simple, such as asking participants about their age, sex and place of residence. However, if your program is reaching specific groups of youth, especially those who are marginalized, it may be more difficult to collect these data. For example, if your program is attempting to reach youth who have been sexually abused, the subject may be too sensitive for participants to respond easily to questions. You may have to ask questions repeatedly and to reassure participants that it is safe to talk.

9. Assessing individual reactions to a program can be difficult.

One measure of quality is how your program is received by stakeholders, staff and youth participants. Assessing how the program is received by these groups will contribute to your understanding of how to overcome social resistance to youth programs. It will also help you determine if your program is headed in the right direction and identify problems in time to correct them. However, eliciting and analyzing individual reactions to programs is difficult to do.

For example, you may want to engage youth and community members to think critically about their needs and to consider how the program could best reach them. Yet, some individuals may have trouble articulating their needs, or their opinions

[5] Bertrand et al., 1994.

may defy what we know about the factors that influence health outcomes. Some community members think it is dangerous to give reproductive health information to youth, and they may want to censor the media in order to produce positive health outcomes among youth. Others may automatically express views that are in line with social norms and values, even if these views do not reflect the true needs of the community. Youth, in particular, may be reluctant to express negative feedback about the program to evaluators, who are often older and carry more authority.

Measuring the social and cultural context of youth development is difficult and may require time and resources that many programs do not have.

Similar tendencies may be found in the reactions of program staff and volunteers. Process evaluations encourage staff to reflect on their work, to see its strengths and weaknesses and to consider alternative strategies. Yet, while most people working with youth are deeply concerned and committed, some have a more ideological approach. They may assume that their strategies are working, even if there is little evidence to suggest that this is true. For example, some staff may insist that increasing access to contraceptive services is the best way to produce results, ignoring the fact that for youth who are abstinent, a more important service may be support in reflecting on and supporting a decision not to have sex. Others may think that their commitment and hard work should pay off in results, and find it demoralizing to discuss how their efforts may be misguided. Staff will need a trusting environment and a supportive process to allow for the kind of reflection in which they can admit that program strategies might need modification.

10. Measuring influences on behaviors that didn't occur is difficult.

Many ARH programs are concerned with preventing unhealthy behaviors and influencing developmental pathways. They are often concerned with measuring events that did not occur because of the program intervention. For example, some programs may aim to delay the onset of sexual activity or prevent unwanted sex. Others may try to prevent early marriage, thus attempting to delay young women's first sexual experience and increase the age at first birth to a time when delivery will be safer. Obviously, measuring the absence of certain behaviors is complex. It requires estimating what level of behavior would have existed had there not been an intervention, then explaining why an intervention caused behaviors not to occur.

11. Measuring behaviors at a variety of developmental levels can be problematic.

Although youth programs are concerned with reaching young people throughout a developmental transition, we are not always sure what outcomes should be expected at specific ages. For example, we may be unsure of what the average age at first sex in our target population is. However, measuring outcomes on sexual behavior can be problematic. Some young people may not have heard about certain sexual behaviors and therefore have problems answering questions about them. This could bias results (e.g., when a girl who has held hands with a boy reports that she has engaged in "sexual activity"). Community

members, and sometimes program staff themselves, may believe it is not appropriate to introduce youth to new topics, such as sexual behavior or illegal behaviors, through a data collection effort.

12. Showing the link between health outcomes and youth development can be complex.

Many programs are increasingly concerned with linking health outcomes to youth development. For example, a program may want to demonstrate that increasing girls' education helps to delay first sex and thus has a positive health outcome. However, what aspects of youth development influence health outcomes may be difficult to predict. We cannot assume that developmental factors would have the same influence on health in different settings, as outcomes are embedded in specific and local contexts, each with their own social and cultural values. Measuring the social and cultural context of youth development is difficult and may require time and resources that many programs do not have.

13. Some changes may not be measurable for a long time, and others may be hard to measure at all.

It may be several years before you can observe changes in the health status of young people, as opposed to the relatively short amount of time it takes to observe such outcomes as changes in levels of knowledge. Moreover, some changes in outcomes may occur long after the program is over; for example, a program that promotes delay of first sex among youth ages 10 to 12 may not be able to observe its results for several years after participants take part in the program. It is therefore important to track trends in such behaviors. For many of the outcomes we are concerned with, we do not know how long it will take to bring about changes. Yet, many youth programs are expected to demonstrate changes in longer-term outcomes in a very short period of time. Some programs define their objectives unrealistically and then falsely conclude that the program did not succeed, when, in fact, more time was required to demonstrate the changes.

Similarly, some program strategies, particularly those that deal with social change, are difficult to measure in numerical or quantifiable terms. For example, measuring complex social processes, such as community mobilization and empowerment, can be difficult because conceptually we are not exactly sure how to define these processes, nor articulate how they are occurring.

Community leaders or other key stakeholders may believe that young people in their communities do not engage in risky behaviors, so they feel there is no need to ask questions.

14. Attributing changes in outcomes to a particular program's strategy and activities is difficult.

How can you conclude that the changes you observe in your target population occurred as a result of your program activities? Measuring changes in outcomes alone is not enough to conclude that the changes occurred as a result of your

program. Other events, like shifting economic or social conditions, could have affected the outcomes you are measuring. There may also have been other program activities directed at your target audience, such as a mass media campaign, going on at the same time. Finally, your program could have attracted participants who were predisposed to the positive outcomes you were trying to encourage. The primary way to determine that an observed change in outcome indicators is attributable to your program is to use a strong study design (see Chapter 5). However, planning and implementing a strong study design requires a high level of resources and skills and may not be feasible for some programs.

15. Some types of evaluation may require funds beyond a youth program's resources.

Outcome and impact evaluations can be costly, especially when measuring numerous outcomes or those that are more difficult to assess. If programs cannot rely on existing data sources, they may need to collect quite a bit of new information about the youth populations they reach. Developing survey instruments, conducting correct sampling procedures and collecting data from individuals can all be expensive. Programs that do not have in-house evaluation expertise may also have the added cost of technical assistance or hiring external evaluators.

13 Tips for Addressing the Challenges of Monitoring and Evaluating Youth Programs

1. Monitor what your system is set up to deliver: programs for youth.
2. Base your program activities, and thus your evaluation effort, on theory.
3. Review what is known about the factors that influence health outcomes.
4. Test and document the elements that contribute to your program's effectiveness.
5. Engage in a genuinely participatory process.
6. Ensure that your data collection effort addresses ethical concerns.
7. Be creative in asking sensitive questions.
8. Define your objectives realistically and provide enough time to measure changes.
9. Use a combined qualitative-quantitative approach.
10. Use monitoring and process evaluation data to support the outcome and impact evaluation.
11. Learn by trial and error.
12. Limit evaluation costs when possible.
13. Build on the advantages of evaluating youth programs.

Thirteen Tips for Addressing the Challenges of Monitoring and Evaluating Youth Programs

Program staff and evaluators around the world are honing their skills and developing creative solutions to the tough challenges of monitoring and evaluation. Below are tips from practitioners in the field and suggestions on how to use this Guide to address measurement challenges.

1. Monitor what your system is set up to deliver: programs for youth.

Monitor the elements of your program's system that respond specifically to the needs of youth. In the Indicator Tables in Chapter 10, we provide some notes on how you can develop a new system or adapt an existing system to capture the needs of youth programs.

2. Base your program activities, and thus your evaluation effort, on theory.

Basing program strategies on theory helps articulate how programs are working and, if they are successful, aids in their replication and adaptation. The Logic Model introduced in Chapter 2 is an example of how a program can plan its activities based on theories of health behavior and social change. Increased understanding of how these and other theories apply in different contexts will strengthen our ability to undertake scientifically sound outcome and impact evaluations.

3. Review what is known about the factors that influence health outcomes.

To help you demonstrate the link between health outcomes and development needs, you should first review what is known about the influences that you assume will affect outcomes. As Chapter 2 suggests, the best way to do this is to review the existing research and literature about your target population. However, if you are unable to access the published literature, or if it is not well-developed in your setting, you can review your staff's experience or talk to colleagues from other organizations. You may also find ways to assess these influences through creative data collection, such as asking questions about a particular behavior in a number of different ways, or modifying language and terminology on your survey instruments to reflect the most important issues in your setting.

Some programs define their objectives unrealistically, and then falsely conclude that the program did not succeed, when, in fact, more time is required to demonstrate the changes.

4. Test and document the elements that contribute to your program's effectiveness.

A number of elements contribute to a youth program's effectiveness. For example, the design elements proposed in the Indicator Tables reflect the current state of knowledge about the design features that are key to program success. We suggest that these elements be tested (i.e., used and evaluated to find out whether they are appropriate and effective) or modified, according to your specific setting and program priorities.

The systems you set up in order to implement a youth program are also key to program success and may be more complex than the systems used to implement other reproductive health service programs. Moreover, the criteria for assessing the quality of youth program system functioning have not been systematically tested. Possible criteria for assessing the quality of your system and its operation are suggested in the Indicator Tables, such as:

- recruitment of staff with appropriate skills,
- components of the training program,
- training program participants who have mastered skills,
- content of reproductive health curricula, and/or
- staff performance.

A participatory process also encourages the community and staff to utilize the information from process evaluations and have a sense of ownership of evaluation results.

Incorporating the elements identified by the international experience of youth programs, lessons from the field of family planning and your own intuition and experience is the best way to establish quality programs and services for young people.

5. Engage in a genuinely participatory process.

Evaluation that engages and involves stakeholders and staff is more likely to produce reactions that are critical and honest than those conducted exclusively by external experts. A participatory process also encourages the community and staff to utilize the information from process evaluations and have a sense of ownership of evaluation results. Giving youth and adults the opportunity to discuss and analyze their concerns, and to suggest and enact solutions, may also increase your program's effectiveness in reaching its objectives.[6] Tips for engaging youth, community members and other stakeholders are provided throughout this Guide.

6. Ensure that your data collection effort addresses ethical concerns.

Professional standards of conduct as well as moral principles and values should be exercised in conducting research and evaluation studies. Ethical reviews are designed to consider and mediate the potential risks and negative consequences to participants as a result of their participation in a study or evaluation. Responding to ethical concerns will improve your relationship with the community and enhance your ability to collect quality data. The more ethical your data collection effort, the more honest and reliable the information you collect, which ensures that your M&E results are valid. Strategies for ethical data collection among young people are discussed in Chapter 7.

[6] The literature on participatory process evaluation is well-developed. One resource specific to the context of young adult reproductive health programs is Shah et al., 1999. *Listening to Young Voices: Facilitating Participatory Appraisals on Reproductive Health with Adolescents.* Washington, DC: CARE International in Zambia and FOCUS on Young Adults.

7. Be creative in asking sensitive questions.

Asking questions of a sensitive nature, while difficult, can be done successfully in many different settings. First, you may need to get support from a broad range of community organizations, to whom you will need to make clear why these questions must be asked. Second, you will need to obtain parental consent, particularly for youth who are legally minors. Third, you can employ "skip patterns" to avoid sexually explicit questions about contraceptive use or other sexual practices if youth have not had sex. Additional tips are provided in the discussion on data collection in Chapter 7.

8. Define your objectives realistically and provide enough time to measure changes.

Programs that define objectives unrealistically may lead to false conclusions. You should budget plenty of time before attempting to measure changes in outcomes, and ensure that your objectives clearly state the outcomes that you expect to produce.

9. Use a combined qualitative-quantitative approach.

Qualitative methods can be used to define social and cultural contexts and develop vocabularies for health education programs, each of which contributes to the formulation of instruments to be used during quantitative surveys. Quantitative methods ensure standardized data collection over time and enable definitive measurement of changes in outcomes that can be generalized to the larger population. They can also be used to show that changes are due to your program activities. Qualitative data can then be used to interpret the findings of quantitative surveys and may reveal program results not discovered through quantitative methods.

Qualitative methods can also be used to assess program goals that are difficult to measure quantitatively, such as empowerment and social change. For example, one qualitative approach asks staff and participants to describe the evolution of the program. Employing this method can help us understand what changes were brought about and why. The results of this approach can then be used to develop a quantitative approach to measure whether those changes are producing the intended outcome in the larger community.

Measuring the absence of certain behaviors is complex.

In Chapter 5, we suggest using a combination of qualitative and quantitative approaches to develop indicators and collect data, which will help you address some of these concerns in more detail.

10. Use monitoring and process evaluation data to support the outcome and impact evaluation.

Conducting outcome and impact evaluations requires resources and time, and even those that are well-designed may not show conclusive results. Using monitoring and process evaluation data can strengthen the results of your outcome and impact evaluations. For example, your MIS may collect information about exposure to the program's services, such as contact with peer educators. If you are trying to demonstrate such outcomes as the increased use of condoms, you may want to measure whether youth received condoms or referrals from a peer educator. While this will not give you conclusive evidence about

your program's effect on the entire target population, it may help you demonstrate associations between positive outcomes and exposure to your program's activities.

While the challenges described in this chapter are many, the task of evaluating youth reproductive health programs can be very gratifying.

The information you collect through monitoring and process evaluation will also help you build the case that the changes were a result of your program, even if an impact evaluation is not feasible. For example, your MIS may show that certain activities were carried out more frequently than others. Your process evaluation may determine that young people liked certain messages better than others and became more involved in the program as a result. It may also document that community leaders' support for specific activities resulted in increased participation or the addition of new activities. Documenting factors such as service utilization, program participation and reaction to program strategies will strengthen the case that your program produced the desired outcomes.

11. Learn by trial and error.

For other measurement concerns, we need to learn by trial and error. For example, we are learning that we can ask questions about sexual behavior, even in settings with very traditional values. Who asks the questions, how we ask the questions, or the place in which we ask the questions may need to be modified in each setting. We also need to be creative about generating valid self-reports of risk behaviors, as we often get the "socially desirable" response rather than an accurate account.

At this stage, many of the suggested measurements in this guide have not yet been tested. We will build on what we know as we collect more evidence and as programs like yours undertake more systematic approaches to monitoring and evaluation.

12. Limit evaluation costs when possible.

While outcome and impact evaluation can be costly, there are ways for programs to limit costs and still produce valid results. For example, an evaluation can examine only those outcomes most important to your program. Measuring outcomes that require less costly data collection methods or utilize already-existing data can also reduce costs. Training and utilizing staff to conduct some parts of the evaluation may be feasible for some programs. A sound sampling strategy can help you limit the amount of data collected while not compromising the validity of your evaluation results.

13. Build on the advantages of evaluating youth programs.

While the challenges described in this chapter are many, the task of evaluating youth reproductive health programs can be very gratifying. There may be other fairly simple ways to avoid challenges in measuring outcomes, including randomly assigning youth in school settings, either individually or by classrooms, and following them. You may find communities where the demand for the program

outpaces the supply. You may also choose to evaluate the impact of selected programs, rather than all programs, or to use a delayed treatment design, which is discussed in Chapter 5. If a few youth are in desperate need of the program, allow them to participate in the program but not the study.

There are many advantages to working with youth. They are interested in learning, and changes in this population can occur relatively quickly. Youth are in a period of great vulnerability, and improving outcomes for them is an investment in our future.

PHOTO: Cathryn Wilcox, JHU/CCP

A FRAMEWORK FOR ARH PROGRAM MONITORING AND EVALUATION

CHAPTER AT A GLANCE

- ➤ Considers the multiple factors that shape adolescence
- ➤ Introduces three major strategies used to improve youth reproductive health
- ➤ Discusses the Logic Model, an approach to designing an effective strategy

Understanding Adolescence and Youth Decision Making

Adolescence is not the same everywhere. The definition of adolescence—and even its existence—has long been a subject of debate. Some argue that adolescence is a period in which children attain physical maturity but are not burdened with adult roles and responsibilities. Adolescence, they say, is a phenomenon of modern, industrial societies.[1] Others theorize that adolescence exists in all cultures at all times, and define adolescence as a life phase that involves the management of sexuality among unmarried individuals, social organization and peer group influence among adolescents, and training in occupational and life skills.[2] A recent modification of the latter definition notes that adolescence is a time of heightened vulnerability for girls and critical capability building for youth (ages 10–19) of both sexes, regardless of their marital and/or childbearing status.[3]

Adolescence is experienced differently in every society, and even within societies there may be vast differences in how some youth experience adolescence as compared to others. To develop program outcomes, objectives and interventions that will have the intended impact, you must first understand the specific context of the youth target population with whom you plan to work.

Sociocultural factors influence how young people experience adolescence, and adolescent sexual behavior reflects a variety of norms and expectations. Particularly where there has been considerable social change in recent decades, young people struggle to balance mixed messages and try to sort out what is best for them.

A broad range of social factors influence young people's reproductive health.
The social factors that influence how young people experience adolescence fall broadly into five categories:

- ➤ The individual characteristics of young people, including their knowledge, attitudes, beliefs, values, motivations and experiences
- ➤ Sexual partners and peers
- ➤ Families and adults in the community

[1] Caldwell, 1998.
[2] Schlegel, 1995.
[3] Mensch et al., 1998.

- ➤ Institutions that support youth and provide opportunities, such as schools, workplaces and religious organizations
- ➤ Communities, through which social expectations about gender norms, sexual behavior, marriage and childbearing, are transmitted

These factors influence how much schooling a young person should receive, what the pattern of courtship and marriage is and when a young person is supposed to take on adult responsibilities, such as work and support for the family. Yet, these factors are also often in conflict with one another. For example, peer norms about the appropriateness of boy-girl relationships may be quite different from those of the family and community. Moreover, each of these factors is constantly changing as the world changes. Understanding and responding to these factors is an important part of developing effective ARH programs.

Note

Research findings

The synthesis of research findings presented here represents more than 350 studies, about 250 of which were undertaken in the United States and about 100 of which were undertaken in Asia, Africa and Latin America and the Caribbean. Each study, which was completed after 1975, had a sample size of more than 100 youth, used scientific criteria and reviewed the antecedents of age at first sex, frequency of sexual activity, number of sexual partners, and condom and contraceptive use. Research identified both protective factors and risk factors. The studies from the United States were synthesized by Doug Kirby of ETR Associates (Kirby, 1999b), and most of those from developing countries were reviewed by Ilene Speizer and Stephanie Mullen of Tulane University (Speizer and Mullen, 1997). Additional results are from papers forthcoming from FOCUS, including "Social Influences on Sexual Behaviors of Youth in Lusaka, Zambia," "Protective Factors Against Risky Sexual Behaviors Among Urban Secondary Students in Peru" and "The Influences of Family and Peer Contexts on the Sexual and Contraceptive Behaviors of Unmarried Youth in Ghana."

Research reveals much about how these factors shape adolescent reproductive decision making.

Researchers are increasingly turning their attention to antecedents, factors that precede and influence how adolescents make decisions about sexual and health behaviors. Antecedents can be positive, a *protective factor,* or negative, a *risk factor.* While research can show the relationship of antecedents to sexual decision making, it is more difficult to identify which antecedents most influence reproductive health outcomes.

Following is a discussion of research findings in each of the five realms of influence.

Individual characteristics

In some cases, young people may calculate or negotiate risks before taking them. They may decide to take risks because they feel invincible, are unaware of consequences and/or want to experiment, or because engaging in risks brings them social status or monetary benefits.

Research has found that the level of knowledge about reproductive health and sex as well as community and family norms and values about reproductive health and sex, influences adolescents' reproductive health decisions. For example, young women in Ghana place a high value on early fertility, which is a risk factor for early pregnancy. Self-efficacy, academic performance and motivation to do well in school appear to protect youth from taking sexual risks. Youth who are actively engaged in learning, who place a high value on helping people and who accept and take responsibility are also less likely to take

sexual risks. Behavioral intentions often shape adolescent risk; for example, young people who intend to avoid STI infection are less likely to take sexual risks. Other related risks have been associated with sexual behaviors among youth. The use of alcohol and drugs, smoking, depression and stress, loneliness and running away from home all enhance sexual risk behaviors among youth. Young people who have been the victims of sexual or physical abuse during childhood or adolescence are often more likely to be at risk.

Biological factors also seem to contribute to adolescent risk behaviors. Early physical development and high testosterone levels increase risk-taking. Age and gender also influence sexual risk; in general, boys are more likely to take sexual risks than girls, as are older youth.

Peers and sexual partners
Researchers have found that if youth believe their friends have sex, smoke or use alcohol or drugs, they are more likely to engage in those behaviors. Power imbalances in a partnership, such as age and income differentials between partners, exchange of money or other goods for sex, and sexual pressure from a partner, also contribute to sexual risks. Conversely, a sense of commitment in a relationship seems to protect young people from undesired health outcomes. There is some evidence that males in same-sex relationships are also more likely to take sexual risks than their heterosexual peers are.

Families
Children of families with lower educational and economic levels have been found to be more likely to be at sexual risk. Families may also enhance risk by devaluing children's education, encouraging early marriage and childbearing or discouraging young people from getting information and services.

However, families can also protect youth from behavioral risks. Living with both parents, having positive family dynamics, feeling supported by parents and other adult family members and experiencing proper supervision by adult family members all seem to protect young people from taking risks. Parental values also influence young people; parents and elders who communicate with young people about their values regarding sex have been found to protect the youth from a variety of risks. Research results are less conclusive about the impact of sexual and reproductive health communication between parents and youth on adolescent decision making.

Institutions
Connections to institutions that support and provide opportunities to youth seem to protect youth from making risky decisions. For example, youth who feel connected to a religious organization are less likely to take risks. School connectedness is also a protective factor, as is successful school performance and a supportive school environment. In contrast, some institutions

Note

Sexual risks
Sexual risks are sexual behaviors that put an individual at risk for unplanned pregnancy, STIs, HIV infection or health problems related to pregnancy and childbearing. Specific sexual risks include:

- too-early initiation of sexual activity,
- sexual intercourse without the use of contraception,
- sexual intercourse without the use of a condom,
- sexual intercourse with more than one partner, and
- sexual intercourse with a partner infected with an STI or HIV.

in the community may promote adolescent risk-taking. The presence of a sex industry and widespread access to entertainment venues such as bars and discotheques may enhance young people's risk-taking.

Some evidence exists that connections to youth organizations also protect youth from risky behaviors. Access to organizations that provide leisure activities, counseling and services for sexually abused adolescents seem to protect youth from sexual risk-taking. Connections with other adults in the community through social institutions, such as neighborhood groups, are also generally found to be protective.

Different strategies are needed to influence the many factors — individual, peers, partners, family, institutions, community — that shape young peoples' behaviors.

Communities

Disorganization or instability in a community often influences youth to take risks. High levels of unemployment and migration, low educational levels, poverty, crime, political instability and war all seem to enhance risk-taking. A lack of programs, health and contraceptive services, and educational and economic opportunities in a community also negatively affect young people's reproductive health decision making.

Some social norms, while not as well-documented by research, also appear to influence youth to make decisions that result in negative reproductive health outcomes. Gender discrimination, community norms that do not value adolescent education, restrictions on girls' mobility and cultural expectations to marry and bear children early in adolescence may negatively impact adolescent reproductive health outcomes.

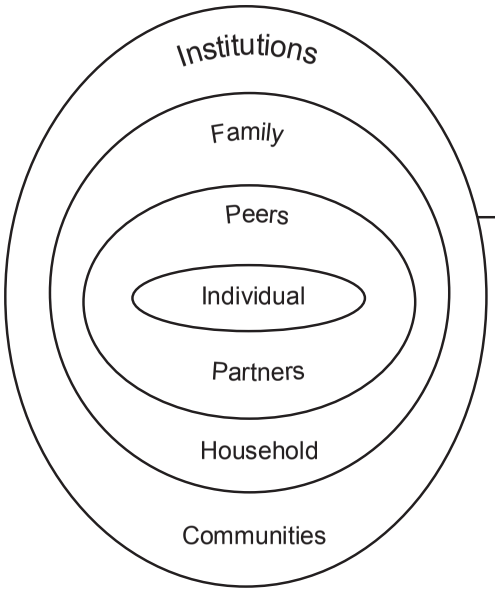
Supportive policies can also protect young people from sexual and reproductive health risks. For example, legalizing contraceptive sales to youth and enforcing a minimum legal age of marriage can be protective actions. Policies that support education and health services for adolescents are also protective. Illegality of abortion and weak enforcement of laws concerning rape and sexual abuse, conversely, may promote negative reproductive health outcomes among youth.

Finally, the mass media influences community norms and values. Advertisements and media that provide positive role models and support responsible behavior can be protective factors. Conversely, exposure to pornography and sexually permissive or violent media may enhance risk-taking among youth.

Three Strategies that Promote Youth Reproductive Health

Globally, programs to prevent adolescent sexual behavior and disease have demonstrated limited results. Two realities largely account for this. First, many of the evaluations have been short-term and are thus unable to show changes in sexual behavior and other reproductive health outcomes, such as pregnancy and STI rates.

Factors that Influence Youth Reproductive Health



Youth Decision-Making and Reproductive Health Behaviors

Reproductive Health Outcomes:

- Fertility
- Abortion
- Morbidity
  - STI/HIV
  - RTI
  - Anemia
- Mortality
- Nutritional Status

Individual

- Age and Gender
- Place of residence
- Knowledge, attitudes and beliefs
- Religiosity
- Self Efficacy
- Skills:
  - motivation to do well in school
  - actively engaged in learning
- Alcohol and drug use
- Other related risky behavior
  - depression, stress
  - running away from home
- Sexual and physical abuse

Peers and Partners

- Perception of peer behaviors
  - perception that peers are sexually active
  - perception that peers are using alcohol or drugs
- Relationship with partner(s)
  - age and income differentials
- Exchange of money or goods for sex
  - sexual pressure
- Sense of commitment with partner

Family and Household

- Low educational and economic levels
- Family attitudes
  - devaluing education
  - supporting early marriage and childbearing
  - discouraging young people's access to information and services
- Harmonious relationship with family
  - quality interaction with family
  - family's values are communicated to youth
  - supervision by adult family members

Institutions

- Connectedness with religious organizations
- Connectedness with schools
  - availability of education
  - a safe school environment
  - academic performance and aspirations
- Availability of youth programs
  - leisure activities
  - counseling
  - sexual abuse services
- Relationship with other adults through community institutions

Communities

- Disorganization (high unemployment levels, poverty, low education levels, political instability, war, crime, high migration)
- Social norms
- Lack of opportunities
- Policy (legality of contraceptives, age of legal marriage, health and education services for youth)
- Policy (illegality of abortion, weak enforcement of rape laws)
- Mass Media (mass media that provides role models and examples of responsible behavior)
- Mass Media (pornography, sexually permissive and violent media)

Second, it seems that the programs most often evaluated—those that provide information about sex and reproductive health, or those that provide reproductive health clinical services to youth—are by themselves insufficient to reduce young people's risky sexual behavior.

Researchers have found that if youth believe their friends have sex, use alcohol or drugs or smoke, they are more likely to engage in those behaviors.

Some studies of adolescent reproductive health programs do, however, suggest directions for future ARH program planning. First, the identification of antecedent risk and protective factors has helped program planners identify and target youth who are at greatest risk of sexual coercion and abuse, unwanted sex, unintended pregnancy, STIs and unsafe childbearing. Evaluations have also found that programs that address a broader spectrum of antecedent influences tend to be more effective at reducing risky behavior or maintaining healthy behavior, and more likely to have a long-term impact.[4] For example, some evaluations have shown that youth development programs that strengthen relationships with school and family result in a reduced age of sexual initiation and lower rates of unwanted pregnancy and STIs. Similarly, programs that develop specific skills related to partner negotiation and condom use have also resulted in desired reproductive health outcomes. Researchers are focusing on other developmental assets to predict and target risk behaviors, such as constructive use of time, presence of a caring community and commitment to learning.[5]

To ensure that they can effectively influence antecedents, programs should also initiate activities based on health promotion, social change and behavior change theories. The three broad strategies described below, when employed simultaneously, can have a maximum impact on young people's reproductive health:

- ➤ Increase knowledge, encourage healthy attitudes, develop skills and form or change youth's behaviors.
- ➤ Improve the social environment so that young people are supported in making healthy decisions and that programs and services are able to operate.
- ➤ Increase access to and utilization of youth programs and health services.

Strategy 1: Increase knowledge, encourage healthy attitudes, develop skills and form or change youth's behaviors.

This strategy aims to influence individual- and interpersonal-level antecedents of adolescent decision making and risk-taking. By focusing on strengthening the individual characteristics of young people, we can help them make healthy decisions about reproductive health. This focus can also influence antecedents at other levels, for instance, by changing community norms, strengthening institutions that support youth and encouraging adults to communicate effectively with young people.

[4]Kirby, 1999c.
[5]Leffert et al., 1998

The transition to adulthood requires specific knowledge and skills.
To make the transition to adulthood, youth need to have the knowledge and skills that help them to:

- participate as citizens (as members of a household, the neighborhood and the larger community, and as workers),[6]
- gain experience in decision making,[7]
- make decisions based on reason,
- assess risks and consequences,
- assess costs and benefits of decisions and actions, and
- interact and communicate with peers, partners and adults.

This knowledge and these skills should be developed from an early age, starting as young as pre-school. They should then be sharpened and strengthened during adolescence in order to make a healthy transition to adulthood.

An increasing body of research indicates that youth development programs that promote the knowledge, skills and other individual assets needed to make a healthy transition to adulthood—coupled with reproductive health information and opportunities to discuss sexuality—can result in a broad array of positive health outcomes.[8, 9]

Youth and adult caregivers need clear and accurate information about sex.
As young people go through physical changes related to human reproduction, they need information and opportunities to discuss sexuality in a safe and open way.

With the influence of global media and changing social values and norms, young people get inadequate, mixed and inaccurate messages about sex. In many societies, talking openly about sex is taboo, so young people resort to friends, movies, videos and pornographic materials for information. Parents and adult caregivers are often uncomfortable discussing sexual topics with their children. Taking the cultural context into account, clear and accurate information should be made available to young people and their adult caregivers through a variety of media and channels.

Note

Sexuality
Sexuality includes not only physical and sexual desires, but also issues of identity, societal and gender roles and human relationships, including those with family, peers and partners.

Young people's attitudes, intentions and motivations to avoid pregnancy and STIs should be strengthened.
Some young people have attitudes, intentions and motivations that encourage them to take sexual risks. For instance, some girls may desire to become pregnant because they think having a baby will bring meaning to their lives or motivate their partners to marry them. Other youth may intend to become sexually active without using condoms because they believe that condoms reduce sexual pleasure. Cultural expectations may encourage young people to marry early and have children soon after. Programs may be able to impact these attitudes, intentions and motivations by providing counseling or small-group discussions for young people, which can help them critically examine their attitudes and change their intentions. For example, a program can help young people examine traditional gender roles and help them make better decisions about what kind of relationship they want to be in, who and when to marry, how much education they want to achieve and how soon they want to

[6] Blum, 1999.
[7] *Ibid.*
[8] Kirby, 1999c.
[9] Leffert et al., 1998.

have children. Programs can also increase young people's motivation to avoid pregnancy and STIs once they closely examine these consequences.

Common Elements of Effective Sexuality Education Programs[10]

- A clear focus on reducing one or more sexual behaviors that lead to unintended pregnancy, STIs or HIV infection
- A foundation in theoretical approaches that have been shown to be effective in influencing other health-related risks
- Ongoing reinforcement of clear messages on risky behaviors
- Basic, accurate information about the risks of unprotected intercourse and methods of avoiding unprotected intercourse
- Activities that address social pressure on sexual behaviors
- Modeling and practice of communication, negotiation and refusal skills
- A variety of teaching methods, designed to involve the participants and have them personalize information
- Incorporation of behavioral goals, teaching methods and materials that are appropriate to the age, sexual experience and culture of the youth
- A duration long enough to complete important activities
- Teachers and peer educators who believe in the program they are implementing

Health education activities can affect many of the factors that influence youth decision making.

A successful ARH program includes activities that influence how young people make decisions as well as the larger environment in which they operate. Some activities common to reproductive health education programs are:

- ➤ sexuality, reproductive health and family life education;
- ➤ skills training, including life skills, vocational skills and skills specific to sexual behavior, such as negotiation and condom use;
- ➤ counseling;
- ➤ peer education and outreach;
- ➤ communications and media outreach; and
- ➤ referrals to health and contraceptive services.

These activities have interacting and overlapping effects; for instance, communications and media outreach may shape community norms about youth, and skills training may stimulate economic opportunities for young people.

Research indicates that the following key programmatic elements of health education activities will lead to improved reproductive health outcomes for youth:[11]

STRATEGY 2: IMPROVE THE SOCIAL ENVIRONMENT SO THAT YOUNG PEOPLE ARE SUPPORTED IN MAKING HEALTHY DECISIONS AND THAT PROGRAMS AND SERVICES ARE ABLE TO OPERATE.

Improving the social environment for youth reproductive health means influencing antecedents that occur among peers, partners, families, institutions and community members. This strategy aims to change social and cultural norms to support young people's healthy decision making, improve programs and policies that reach youth and support adults and institutions that interact with and support youth.

A positive social environment supports healthy lifestyles.

Relationships with friends, partners and family members, as well as the influence of community, school and other institutions, all play a role in shaping multiple health outcomes. Some programs aim to improve the social environment for ARH. This includes encouraging critical discussion of the social and cultural norms that may adversely impact ARH, such as norms

[10] Kirby, 1999b. This table is based on the analysis of evaluations conducted of sexuality education programs in the United States, and may be more or less relevant in some developing country settings.

[11] Kirby, 1997; Choi and Coates, 1994; McKaig et al., 1996; and Houvras and Kendall, 1997.

related to gender roles. Other programs might attempt to strengthen institutions that reach and support youth, such as youth clubs and religious organizations, or develop policies and programs that provide the services youth need.

Supportive and caring communities can make a difference. For example, community organizing builds communities and institutions in ways that enable members to identify and solve problems and respond to needs. It fosters ownership and participation, and engages community members—adults and youth—in social action that considers and addresses young people's reproductive health needs.

Family support plays a critical role in young people's decision making. Parents and other adult family members play a critical role in shaping young people's aspirations and values. Even when adult caregivers have difficulty discussing sex and reproductive health with youth, support from adults can positively influence a young person's reproductive health outcomes. Adult caregivers need to be encouraged to value the education of youth, provide supervision and support and communicate effectively with young people. Programs that reach parents might aim to help parents create a harmonious relationship with their children by practicing what they could say to effectively show support. Programs might also raise awareness among adult caregivers of how some cultural traditions, such as early marriage, have a detrimental effect on young people's lives.

Programs must identify and address the dynamics of youth's social systems. Understanding and addressing these dynamics is also crucial to improving the environment for youth. Many programs work to improve our understanding of social systems and to strengthen and make more responsive those systems that support youth. For example, a program may find that some young people are at a disadvantage—both to adults and to other youth—due to differences in age and experience, gender, income and education. An adolescent reproductive health program may not only try to improve the knowledge and skills of those youth, but also attempt to influence the behavior of those holding power over them.

Strategies to Create a Supportive Environment for Youth

- Mobilize community action, particularly among youth.
- Generate collaborative responses to ARH among youth, community members, and institutions and organizations working in the community.
- Raise awareness of young peoples' needs and the social, cultural, economic and political issues that contribute to their RH concerns.
- Conduct mass media and social marketing campaigns.
- Gain stakeholder and other adult support for discussions with, and activities and services for, young people.
- Address antecedents that contribute to youth RH risks, such as dropping out of school, gender inequity, early marriage, female genital cutting, the sex industry and drug and alcohol consumption.
- Improve other sectors in related areas, such as female education and vocational training.
- Overcome resistance to providing RH information and services to young people, and ensure that these services are affordable.
- Institute policies to promote access to reproductive health information, education and services. Remove restrictions that limit this access.
- Support networks and coalitions to encourage advocacy, service referrals and broader social changes.

Strategy 3: Increase access to and utilization of youth programs and health services. This strategy focuses on providing the opportunities, programs and services that allow young people to gain access to youth programs and health services. By strengthening the institutions that support youth, such as youth clubs, recreational facilities, religious organizations, schools and health facilities, this strategy aims to

influence individuals' participation. The existence of youth programs may also influence families, institutions and communities as they increase the visibility of youth engaged in positive activities and change adults' attitudes toward them.

Youth programs can affect young people's lives on multiple levels.
Many youth programs aim to increase the number of young people who participate in activities that build their skills, build positive relationships with peers and adults and provide a creative outlet for their energy. For example, youth programs may attempt to build young people's skills; encourage activism in the community; provide sports, arts or other creative activities; or foster adult mentoring of youth. At the individual level, these programs help to build self-esteem and skills and encourage young people to have aspirations for the future. At the interpersonal level, they encourage the creation of healthy norms among peer groups and positive interaction between young people and adults. At the community and institutional levels, youth can have a direct influence on changing the environment if encouraged to participate as advocates for youth-related programs and policies. The presence of youth organizations can also influence how adults in the community view youth and help the community see young people as an important asset.

Families can protect youth from behavioral risks.

Connectedness to schools improves young people's knowledge and skills.
Increasing the quality and quantity of education young people receive is another strategy to improve adolescent reproductive health. In places where young people have few educational opportunities, increasing local and national commitment to education can be an important part of addressing adolescent reproductive health. In addition to improving access to education, schools can improve their physical and emotional environments. Programs can address sexual harassment in schools, change school policies that do not allow attendance by pregnant adolescents, improve safety in schools or strengthen the extracurricular activities.

Religiosity and connection to religious organizations can positively influence youth.
Adolescence is a time of rapid change, and religious beliefs can help young people understand and process the challenges they face. Religiosity may have more to do with a young person's strong religious belief than it does with his or her actual attendance or participation in religious activities. In addition, evidence shows that feeling connected to a religious organization can support young people in making healthy decisions. Programs may want to increase these links while respecting individual decisions in this realm.

Health services enable young people to act on their healthy decisions.
The provision of health services, such as counseling, contraceptives, maternal care and nutrition programs, to youth is crucial. Without such services, young people may not be able to act on the positive decisions they make. In order to reach more young people with health services, we need to understand how young people prevent reproductive health problems and seek

treatment both within and outside the formal service delivery systems.

Many would agree that, in order to make healthy decisions about illness, it is important to see a trained medical service provider. Yet, reproductive health programs have largely addressed older, married women and sometimes men; young people perceive—often correctly—that family planning and STI clinics would not welcome them. The barriers to youth access of health services are numerous:

- ➤ long distances to service locations, and unsafe or unavailable transportation;
- ➤ inconvenient hours of operation;
- ➤ lack of anonymity;
- ➤ concerns about privacy and confidentiality;
- ➤ staff attitudes and actions, including scolding and moralizing;
- ➤ fear and embarrassment;
- ➤ cost of services; and
- ➤ laws and policies that make serving youth difficult.

Many youth rely on resources outside the formal health service provision system. These resources may include home remedies, traditional methods of contraception and abortifacients, provision of contraceptives through friends or relatives, clandestine abortion, and contraception and medication purchased without a doctor's prescription from pharmacies or traditional health practitioners. Many programs are trying to increase young people's utilization of reproductive health services through activities and strategies that:

Characteristics of "Youth-Friendly" Health Services

Health Provider Characteristics

- Staff specially trained to work with youth
- Respect for young people
- Privacy and confidentiality honored
- Adequate time for interaction between client and provider
- Peer counselors available

Health Facility Characteristics

- Separate space and special times set aside
- Convenient hours
- Convenient location
- Adequate space and sufficient privacy
- Comfortable surroundings

Program Design Characteristics

- Youth involvement in design and continuing feedback
- Drop-in clients welcomed and appointments arranged rapidly
- No overcrowding and short waiting times
- Affordable fees
- Publicity and recruitment that inform and reassure youth
- Boys and young men welcomed and served
- Wide range of services available
- Necessary referrals available

Other Positive Characteristics

- Educational material available on site, which can be taken home
- Group discussions available
- Possible to delay pelvic examinations and blood tests before receiving contraceptives
- Alternative ways to access information, counseling and services outside of a formal health facility

- ➤ increase young people's knowledge about the availability of reproductive health services;
- ➤ generate demand for services, for example, by promoting services through peer outreach workers; and
- ➤ examine where and how young people seek information and treatment, and improve the "youth-friendliness" of those services.

The following items are seen as characteristics of effective "youth-friendly" health services, whether services are provided in a clinic, hospital, pharmacy, youth service organization or other venue:[12]

Identifying Appropriate Program Activities

We now have a clearer understanding of the multiple levels of influence on adolescence and the broad strategies that promote healthy reproductive behavior among youth. Using this understanding as our foundation, we can design programs that are more likely to be effective and, thus, worth the effort of good evaluation. These programs will:

- ➤ clearly define desired health outcomes,
- ➤ identify the protective and risk-enhancing antecedents that influence those outcomes, and
- ➤ use program strategies that respond to more than one of the antecedents that impact adolescent reproductive health outcomes.

Note

The Logic Model

The concept of a "Logic Model," and its importance to the design and evaluation of youth programs, was introduced by Kirby during a presentation to a meeting on Adolescent Health and Development, Washington, D.C., 4–6 February 1999. It is a simplified version of the logical framework, which emphasizes that outcomes should be pursued based on antecedents identified by research.

One way to design your strategy while keeping these elements in mind is to use a *Logic Model*.

The steps, outlined below, are as follows:

- ➤ Define your program's goals and desired behavioral outcomes (the process of defining goals and outcomes is discussed in detail in Chapter 3).
- ➤ Identify the antecedents that, according to research, influence—both positively and negatively—the behavioral outcomes your program desires. In many places, there may not be enough research to suggest the whole range of factors that influence youth behavior and decision making. In this case, you may base your assumption of influences either on a review of the research suggesting antecedents in other countries, or you can use your experience with youth to make a "best guess" about the antecedents that influence health outcomes. You should also try to directly ask youth what they think influences their decision making.
- ➤ Identify one or more program activities that you think, based on your own experience or on the international literature about what works, will specifically influence each antecedent.

[12] Senderowitz, 1999.

Define Your Program's Goal	Define Your Program's Desired Behavioral Outcomes	Identify the Antecedents of Behavioral Outcomes Your Program Desires	Identify Program Activities that You Think Will Influence Each Antecedent
• Decrease rates of pregnancy and STIs among youth ages 14–19 in our district.	• Decrease premarital sex • Increase use of condoms among sexually active youth	• Community norms about premarital sex and appropriate age of sexual initiation	• Develop education program to encourage adults to discuss norms around premarital sex with youth
	• Increase age of sexual initiation • Increase age of marriage	• Opportunities for education	• Initiate community mobilization campaign to change norms that do not value girls' education • Lobby for expansion of opportunities for secondary education
		• Individual's ability to say "no" to sex	• Provide life skills education emphasizing how to say "no" to sex in school health education program
		• Individual's ability to use contraception	• Establish peer education program to reach sexually active youth
		• Youth's access to condoms, contraception and clinical services in a confidential way	• Encourage development of national health policies that support provision of services to youth • Include youth representatives in clinic advisory committee • Establish "youth-friendly" and confidential services (e.g., educating health workers) at both national and local level
		• Community norms about appropriate age of marriage	• Influence community norms to support later age at marriage

For example, by using the Logic Model to guide program design, your assumptions about what influences your desired outcomes are both clear and specific to your context. You are more likely to think broadly about the factors that influence your desired behavioral outcomes and to include contextual influences. Finally, your program activities will be designed to link directly to the influences you think will affect the behaviors you are concerned with.

Key Elements of Successful ARH Programs[13]

- **Strategic planning**: Effective programs clearly state process and behavioral objectives before the program begins, as a prerequisite to measuring success.
- **Target audience identification**: Young people have diverse needs, depending on their varied characteristics, such as age, school status, marital status, gender, family characteristics and experience. In designing programs, it is important to identify the specific target group and address its needs accordingly.
- **Needs assessment**: Understanding the specific issues and needs of the youth who are expected to participate in or receive services from the program ensures that the program's design and content are appropriately shaped.
- **Youth involvement**: Youth are best able to identify their own needs and will feel more ownership of a program when they are included in design and implementation.
- **Community involvement**: Community members, such as policymakers, health professionals and religious leaders, should be involved in program planning to ensure support and acceptance.
- **Adult involvement**: Involving parents and other adult family members may help to ensure that the program does not meet with resistance and to educate parents about reproductive health issues and adolescent needs.
- **Protocols, guidelines and standards**: Specific and detailed operational policies governing how a program should serve youth may help to encourage a consistent level of quality, particularly when service providers are unfamiliar with the youth population.
- **Selection, training and deployment of staff**: Staff providing services to young people require specific qualities, training and supervision to ensure that clients are well-treated and to ensure client retention.
- **Monitoring and evaluation**: Collecting data helps managers monitor performance, evaluate outcomes and impact and improve program strategies.

[13] Israel and Nagano, 1997; Birdthistle and Vince-Whitman, 1997; Senderowitz, 1997a; and Senderowitz, 1997b.

Learning from the International Experience with Youth Reproductive Health Programming

The design of ARH programs can be informed by the experience and evaluation of programs from around the world. The "key elements" of ARH programs listed below were compiled through a literature review. While these elements have not been systematically tested in field settings, we offer them here because they may help in your program design and implementation.

Note

Additional information on key elements
The key elements of adolescent reproductive health programs are presented in four papers produced by FOCUS on Young Adults. These papers can be accessed and downloaded from the FOCUS Web site at <www.pathfind.org/focus.htm>.

PHOTO: Marsha McCoskrie, JHU/CCP

DEVELOPING AN ARH MONITORING AND EVALUATION PLAN

CHAPTER AT A GLANCE

- Defines program goals, outcomes and objectives
- Helps you define the scope of your ARH monitoring and evaluation effort
- Offers guidance on how to plan and conduct a monitoring and evaluation effort, using the rest of this Guide

Establishing Goals, Outcomes and Objectives for Youth Reproductive Health Programs

This section discusses the goals, program outcomes and objectives of an adolescent reproductive health program, which form the basis of your M&E effort. Each is a different expression of the reproductive health outcomes the program is trying to achieve.

Goals define the overall impact your program hopes to have.

A goal states the impact a program intends to have on a *target population.* The target population is the specific group of individuals your program is trying to affect, and can include youth as well as the adult service providers, teachers, family members or community members who interact with young people. ARH programs often have the general goal of improving the reproductive health of young people. Goals may be stated more specifically depending on the reproductive health needs of the youth population.

Note

Terminology

People working in evaluation use many different terms to describe what they do. The existing evaluation terminology is often interpreted differently in different settings, and sometimes evaluators spend too much time debating which term is best to use. In this Guide, we use terms and concepts that are intended to reflect the stages and components of youth programs as they are implemented in the field. We have defined them in ways that we hope will be understandable and accessible to those who do not have a research background.

Program outcomes are the specific results that your program hopes to achieve.

Your program's intended outcomes are related to your established goals, such as a decrease in STI rates or improvement in nutritional status. To produce these outcomes, programs focus on intermediate behavioral changes, such as the delay of sexual initiation, increased use of condoms or contraception or increased breastfeeding. Programs can establish short-term, intermediate and long-term program outcomes, as detailed on the next page.

Goal: To Improve Young Women' s Reproductive Health		
Short-Term Outcome	**Intermediate Outcome**	**Long-Term Outcome**
Improve the quality of interactions between parents and youth ages 10– 19 in our district.	Increase the average age at sexual initiation among youth ages 14– 19 in our district.	Decrease pregnancy rates among youth ages 14– 19 in our district.

Objectives are explicit, measurable statements of program outcomes.

There are two kinds of objectives: *population-level* and *program-level.*

Population-level objectives state intended results in terms of the target population and are directly related to the outcomes identified by your program. They describe what impact your program hopes to have in the youth population it aspires to reach, influence or serve. For example:

- ➤ Increase the average age at sexual initiation among youth ages 14–19 in our district by one year.
- ➤ Increase the percentage of youth ages 14–19 in our district who are actively involved in youth organizations that provide leisure activities.

Program-level objectives state intended results in terms of the structure, management or operations of a program. They describe the activities you will undertake to achieve the impact your program hopes to have. For example:

- ➤ Train 30 peer educators to provide quality counseling to youth every six months.

Measuring Objectives

How you conceptualize and express your objectives and their measures will frame your actions.

The measure of an objective should be stated in terms of targets.

Targets are the level of the objective you plan to achieve within a stated time.[1] Targets may be either *quantitative* (numeric) or *qualitative* (descriptive), depending on the nature of the activity and the indicator chosen to measure it.[2] Targets may express quantity (how much), quality (how well) or efficiency (least cost per outcome produced).

The target of population-level objectives should be defined by referring to baseline information.

Baseline information describes the current status or situation in a community before an intervention takes place. Baseline information is important because it provides points of comparison against which you will measure whether your objectives were accomplished. If baseline information is not available, you may need to collect information about the target population and its needs before your program begins. This will provide you with starting measures that can be the basis for an outcome or impact evaluation that the program undertakes later.

[1] Targets are quantitative estimates that are used for the purpose of budgeting, planning and tracking changes in outcomes. They should not be understood as quotas, or used as a basis to coerce any individual to accept services, such as contraception, that are inconsistent with his or her moral, philosophical or religious beliefs. Targets should not be used as a basis for compensation of service providers. All youth reproductive health programs should, of course, safeguard the rights, health and welfare of all individuals who take part in the program

[2] Indicators are discussed in detail in Chapter 4.

The source of baseline information could be:

- ➤ a survey of youth prior to the intervention;
- ➤ data documenting prior youth program experience;
- ➤ external measures collected by another organization, government agency or donor, such as government health facility utilization data;
- ➤ information on youth reproductive health obtained from a national survey, such as a Demographic and Health Survey (DHS); or
- ➤ the professional judgment of those who work with youth.

For example, your prior program experience may tell you that only 5 percent of youth are seeking counseling services from peer educators in the schools where your program functions. However, you are aware that a partner organization in a neighboring district found that 8 percent of youth sought counseling. Referring to this baseline information, you might determine that your program objective should be "Increase the percentage of youth ages 10–19 who seek counseling services from peer educators to 10 percent within one year."

- ➤ The target of program-level objectives should be defined by program experience.

To determine targets of program-level objectives, such as numbers of peer educators who should be trained, refer to program experience and resources. For example, you might determine that to reach the 500 youth in your target population, you would ultimately like to train 40 peer educators. Since your budget only allows for one training every six months, and experience has shown that training 20 peer educators at a time is most effective, you might want to set your target as training 20 peer educators every six months.

Monitoring and evaluation requires an understanding of measurement and indicators.

Measurement is the use of methods and procedures for systematic observation and assessment.[3] A variety of methods and procedures are used to collect information about your program and its target population.[4]

To measure how a program is functioning and what outcomes it is having in the target population, you will use *indicators.* An indicator is a measure of program objectives and activities.[5] Changes in indicators demonstrate that a program is functioning and the effect—positive or negative—it is having on the target population.

Information is collected on some of your objectives—both program-level and population-level—in order to measure whether a program's activities are being implemented, the quality of program implementation, to what extent the program is being utilized, or the changes that are taking place in your target population, if any. In general, information collected during a process evaluation will measure program-level objectives. Information collected during an outcome or impact evaluation will measure population-level objectives. To measure changes in objectives, baseline information is compared to data collected after the program has been operating for some period of time.

[3] Green and Lewis, 1986.

[4] Data collection is discussed in detail in Chapter 7.

[5] Indicators are discussed in detail in Chapter 4.

How to Translate Goals and Outcomes Into Objectives		
Goal: • Improve the reproductive health of youth ages 10– 19 in our district.	To translate to outcomes, describe the specific results your program hopes to achieve.	**Outcomes:** • Increase condom use among sexually active youth ages 15– 19 in our district. • Increase the number of sexually active youth under the age of 19 who discuss condoms with their partners.
Outcomes: • Increase condom use among sexually active youth ages 15– 19 in our district. • Increase the number of sexually active youth under the age of 19 who discuss condoms with their partners.	To translate outcomes into population objectives, refer to baseline data: The most recent DHS found that 5 percent of youth ages 15– 19 in your district use condoms at first intercourse. A baseline survey conducted by your organization found that 15 percent of sexually active youth ages 15– 19 have ever discussed condoms with their current sexual partners.	**Population-level objectives:** • Increase condom use at first intercourse by youth ages 15 – 19 in your district to 10 percent within one year. • Increase the number of sexually active youth under the age of 19 in your district who have ever discussed condoms with their current sexual partner to 25 percent within one year.
Outcomes: • Increase condom use among sexually active youth ages 15– 19 in our district. • Increase the number of sexually active youth under the age of 19 who discuss condom use with their partners.	To translate outcomes into program-level objectives, describe the activities you will undertake to achieve the outcomes.	**Program-level objectives:** • Train 25 peer educators to hold skills-building sessions with youth ages 15– 19 about condom use and negotiation. • Hold 30 skills-building sessions with youth ages 15– 19 about condom use and negotiation.

Monitoring, process evaluation, outcome evaluation and impact evaluation involve the following steps:

- ➤ agreeing on the scope and objectives of your M&E plan with stakeholders,
- ➤ selecting indicators,
- ➤ systematically and consistently collecting information on those indicators,
- ➤ analyzing the information gathered,
- ➤ comparing the results with the program's initial goals and objectives, and
- ➤ sharing results with stakeholders, including youth.

A strong M&E plan should use indicators to measure both population-level and program-level objectives.

Population-level objectives relate most directly to the sexual and reproductive health outcomes your program hopes to achieve. However, they are often difficult to measure because they deal with sensitive issues, such as whether or not young people are having sex. You should always try to measure population-level objectives

related to intermediate behavioral outcomes, but you may have difficulty doing so.

Measuring short-term objectives related to the risk and protective factors your program thinks influence young people's behavior is important for two reasons. First, in the absence of showing changes in behavior, the achievement of short-term objectives is a good sign that your program is producing outcomes. Second, measuring short-term objectives also helps test your assumptions about the factors that influence the behavior and decision making of young people. This information may provide insights into how your program strategy is working, or not working, to influence the behavior that produces the long-term reproductive health outcomes you are concerned with.

Measuring program-level objectives is an important part of understanding how your program is working. Program-level objectives are measured during a process evaluation, and provide information on how a program is functioning. A process evaluation may offer insights into why your program is having an impact (or not) and is important if you plan to scale up or replicate the strategy your program uses.

Defining the Scope of an M&E Effort

Scope refers to the extent of the activity you will undertake in a monitoring and evaluation effort. The scope of your M&E effort is determined by several factors. Ask yourself six key questions:

- ➤ What should be monitored and evaluated?
- ➤ When should ARH programs be monitored and evaluated?
- ➤ How much will M&E cost?
- ➤ Who should be involved in M&E?
- ➤ Who should carry out the evaluation?
- ➤ Where should M&E take place?

Each is discussed below.

What Should Be Monitored and Evaluated?

M&E can measure each stage of your program's development: design, systems development and functioning, and implementation. After you have developed goals, objectives and activities, your next step is to make decisions about M&E in each of these stages. Your M&E effort can measure each stage to determine how the program is working and its impact on the target population. You can review each stage for ideas and options for M&E efforts.

Program design is measured by process evaluation.

A community needs assessment often forms the basis for program design. The process of program design involves developing a strategy or systematic approach to address the community's needs, identifying actions and activities required to implement the strategy, and identifying the resources needed to carry out the activities. Assessing how well a program has been designed is one aspect of process evaluation because the program design affects the success of a program. Documenting the problems with program design will help explain why a program did not achieve its objectives; conversely, if a program is successful, documenting will help explain what key design elements contributed to its success. Those elements can then be used to expand or replicate a program. Chapter 5 includes information on how to monitor and evaluate the design stage of a program.

Worksheet 3.1 Identifying Program Goals, Outcomes, Context and Objectives	
1. What are the program' s goals?	
2. What short-term, intermediate and long-term outcomes does your program hope to achieve?	
3. What short-term, population-level objectives does your program hope to achieve (including objectives related to antecedent factors)?	
4. What intermediate, population-level objectives does your program hope to achieve?	
5. What are the program-level objectives? How will you achieve the population-level objectives stated above?	
6. What activities will be implemented by the program?	
7. Who are the stakeholders of the program?	
8. How might the local political or cultural context affect the program?	
9. Will current economic conditions affect program implementation or participation by youth?	

Systems development and functioning is measured through monitoring and process evaluation.

Systems development involves the creation of a management and support system to carry out the program. Support systems include MIS, financial management systems, personnel systems, and commodities and logistics systems. Conducting preparatory activities such as recruiting and training staff, developing curricula, drafting service guidelines and developing IEC or behavior change communication (BCC) materials is an important part of systems development.

Systems functioning involves the ongoing performance of the systems used to operate the program and includes issues such as how decisions are made within the program, whether internal and external communication channels are functioning well, how well coordination between regional programs and headquarters is conducted, whether training and supervision are ensuring quality performance, and personnel job descriptions and job performance.

If you are able to document how a program's systems are functioning, this will help explain why a program is—or is not—working. To determine how a program's systems are functioning, monitoring and process evaluation should:

- document the development of support systems and determine if they are actually operating once program implementation begins;
- assess the performance of support systems; and
- measure how effective the preparatory activities are in readying program personnel for program implementation.

Implementation is measured through monitoring, process evaluation and outcome/impact evaluation.

Implementation is the process of carrying out program activities with the target population and providing services to them, i.e., the actual performance of your planned activities. For example, the activities of a youth center may include hiring and training staff and volunteers, holding educational sessions at the center, involving youth in developing leisure activities and providing counseling services to young people.

Documenting the problems with program design will help explain why a program did not achieve its objectives and—if successful—what key design elements contributed to the success.

Monitoring and process evaluation reveal how program implementation is occurring. Outcome and impact evaluation help determine whether your program is achieving its objectives by measuring changes in outcomes in your target population. Together, this information should help you explain why the program is—or is not—reaching its objectives, and contributes to an understanding of program outcomes.

Activities	Monitoring & Process Evaluation	Outcome & Impact Evaluation
Design Stage • Determine whether peer educators are an effective way to reach your target population.	• Were youth in the target population consulted about whether they thought peer educators would be effective?	N/A
Systems Development & Functioning Stage • Develop curricula to train peer educators. • Recruit, select and train peer educators.	• How many peer educators are recruited, selected and trained? • What is the quality of the training provided to peer educators?	N/A
Implementation Stage • Peer educators provide counseling three afternoons a week in five health clinics.	• How many youth are counseled by peer educators? • What is the quality of the counselin provided by peer educators?	• Do changes in knowledge, attitudes and behavior occur among youth who are counseled by peer educators?

The goals, objectives and activities of your program shape the scope of what will be monitored and evaluated.

By identifying every activity your program has undertaken at the design, systems and implementation stages, you define the scope of your M&E effort. At each stage, your activities should be monitored and/or evaluated. The table on the next page illustrates how activities undertaken at each stage of a peer education program might be monitored and evaluated. Identifying your activities at each stage and defining the possibilities for M&E is the first step in determining scope.

How you plan to use M&E information shapes what will be monitored and evaluated.

Your intended use of M&E information will help you determine the scope of your M&E effort. Possible uses include the following:

- **Ensuring that program activities are carried out as planned:** If so, you will need to track changes in program-level objectives through an effective monitoring system.
- **Assessing how well activities are being carried out and making improvements as needed during the course of program implementation:** If so, you should undertake a process evaluation.
- **Determining whether changes in outcome indicators are occurring in the target population for your program:** If so, you should conduct an outcome evaluation. If you have more resources and are interested in showing how much of the observed change in outcome indicators is due to your program, then you should conduct an impact evaluation.

Reasons to Monitor and Evaluate: Different Needs for Different Stakeholders

By developing consensus among stakeholders about what information should be collected, given your available resources, you can make an M&E effort more manageable.

Program Managers and Staff	Funding Agencies and Policymakers	Communities and Youth
What M&E Measures: • Quality of activities and/or services • Why some sites are less successful • Capacity in M&E techniques • Program coverage	**What M&E Measures:** • Evidence of achievement of program objectives • Program outcomes and impact • Program cost-efficiency • Data about youth reproductive health	**What M&E Measures:** • Youth behaviors related to reproductive health • Young people' s needs • How program funds are being spent • The process and impact of community participation
What M&E Results Identify: • Priorities for strategic planning • Training and supervision needs • How to improve reporting to funding agency • Feedback from clients • Why program is not accomplishing what it set out to do	**What M&E Results Identify:** • Priorities for strategic program funding • Programs that qualify for donor assistance • Best practices that donor should require of youth programs • Impact of donor assistance	**What M&E Results Identify:** • Actual and potential benefits of youth programs • Need for new and better youth services • Community resources that can be used to support ARH programs • Need for local support for ARH issues and action
What Decisions Are Guided by M&E Results: • Resource allocation • Replication and scaling up of interventions • Fund-raising • Motivating staff • Policy advocacy • Community mobilization	**What Decisions Are Guided by M&E Results:** • How much funding should be allocated to ARH • What types of youth programs should be funded • Which program approaches should be presented as models • New strategic objectives, activities or results packages • Replication and scaling up of successful programs	**What Decisions Are Guided by M&E Results:** • The degree to which community members and youth should participate in and support the program • How to better coordinate community actions to address ARH • How many and what type of local resources should be allocated to ARH

- **Responding to donors' requirements:** Some donors may require programs to undertake outcome or impact evaluations.
- **Understanding how your program is performing and what outcomes it is influencing:** This will help you decide whether to continue, change or expand your program strategy.

When Should ARH Programs Be Monitored and Evaluated?

Monitoring and process evaluation should occur throughout the life of a program. The information you collect can be used to ensure that you are meeting objectives, to improve program performance and to provide feedback and support to staff and program participants.

If You Start M&E Too Late . . .

- you may not have baseline information on the status of your target population before your program began;
- the information you collect will be less meaningful;
- the information will not be useful to make improvements in program strategy; and
- your evaluation results will be less conclusive about whether changes in outcomes occurred, or whether changes can be attributed to your program's activities.

Outcome and impact evaluations are usually done near the end of a program, although they often use baseline information gathered at the program's start. An impact evaluation has to be included in a program's design from the beginning or you will not have the type of baseline information needed to measure changes in outcomes and then attribute them to your program.

It is very important not to conduct an outcome or impact evaluation prematurely. For some intended outcomes, such as changes in risk behaviors, program activities need to be carried out for some time, perhaps several years, before changes in the target population can be observed. In this case, outcome or impact evaluation may take place after the program has been fully functioning for some time.

When to conduct evaluations should be based on your program's objectives, the needs of various stakeholders for information about the program, your knowledge of the program, available resources and your judgment as a manager. The point in your program at which you start an M&E effort will determine the type of monitoring and evaluation you can undertake.

Starting M&E at the beginning of a program is ideal.

Monitoring and evaluation should be planned—and started—at the beginning of any new program. Early planning allows you to define your M&E effort based on your objectives and activities, and to be strategic about what you plan to measure. It also enables you to find existing information and collect baseline information at the ideal time—your program's starting point. This will allow you to conduct either outcome or impact evaluations with greater ease and enhances your ability to measure the program's true impact. Starting monitoring and process evaluation early also allows you to use M&E results to make improvements in the program as it is being implemented. Finally, starting early allows you to ensure that M&E costs are adequately covered by your budget.

Flow of an M&E Effort Started at the Beginning of a Program			
Stage of Program	**Monitoring**	**Process**	**Outcome/Impact**
Early	Set up monitoring system (MIS); identify indicators and instruments; plan for tracking program, data analysis and reporting.	Assess systems development and functioning, including training and supervision of staff. Provide early feedback. Assess if program is responsive to youth or if it needs any additions.	Identify objectives and indicators. Take baseline measurements. Create an outcome or impact evaluation plan.
Middle	Assess MIS and data. Modify if original system is inadequate or if program adds new components. If program is not performing as planned, launch process evaluation.	Conduct more formal mid-term process evaluation to assess quality of program performance. Determine coverage, or whether the program is reaching its intended audience.	Take mid-term measurements. Analyze short-term outcome measures, such as changes in knowledge, increase in use of programs and changes in contextual factors. Provide feedback to program.
Late	Analyze data from tracking system to conclude if you conducted the program as planned. Prepare and submit reports.	Analyze end-of-program measurements. Determine what was done to improve quality of program' s implementation. Make recommendations for program replication or expansion.	Take end-of-program (follow-up) measurements. Examine evidence of changes in outcomes. Depending on study design, conduct impact analysis to conclude whether outcomes are attributable to program activities. Report to donors and other stakeholders.

Some activities can still be measured if M&E is started in the middle of a program.

You may realize that you need an M&E plan later—after the program has started. If you start your M&E effort in the middle of your program, its scope will probably be limited. It may still be possible to conduct an outcome evaluation, but you will probably have to use baseline information taken after the program's start. While the results may not be as clear and strong, they may still be useful. While an MIS can be set up mid-program to track monitoring and process evaluation results, it will be less useful than one launched at the beginning.

Even fewer activities can be measured if M&E is started toward the end of a program.

Some program managers may not think about what they are going to monitor or evaluate until the program is almost complete. If you start your M&E effort at the end of your program, your options are severely limited. First, it is of little use to set up a monitoring system at the end of a program. While you can assess program activities in retrospect (by soliciting participant and stakeholder feedback after the program is well underway), you may produce biased results. Finally, while an outcome evaluation is possible, it will have to rely on external standards—estimates of the plausible status in your community before the intervention took place—as comparison data. These standards may or

may not accurately reflect the knowledge, attitudes and behaviors of your target population before the program began, thereby limiting your ability to demonstrate change in outcomes.

How much will M&E cost?

Your financial resources will influence the level of evaluation you take on. Program managers must determine whether the time, effort and cost of an evaluation are justified in light of the expected benefits. If you have no staff capable of conducting an evaluation—or cannot release trained staff from other duties to concentrate on doing sound M&E—and if you cannot afford to hire an outside evaluator, you may elect to carry out only a very basic review of your program's progress.

An impact evaluation has to be designed into a program from the beginning or you will not have the type of baseline information needed to measure changes in outcomes and attribute them to your program.

If you have few resources, your first priority should be to establish a monitoring system.

The best use of limited resources is to establish an effective monitoring system, so that you can ensure and document that your program was implemented according to plan.

If additional resources are available, undertake some form of process evaluation.

Some types of process evaluation can be done quite inexpensively, e.g., by having supervisors periodically observe service delivery or interview program clients as part of their duties. More systematic process evaluations (such as conducting focus groups with youth) require more resources.

Outcome evaluations require a moderate to high level of resources.

You will need to decide early if you are going to do an outcome evaluation so that you can budget accordingly. The cost will largely depend on how many outcomes you want to measure and the level of difficulty involved in measuring them. It will also depend on what data sources already exist and how much new data you will need to collect. The following steps can help you contain the costs of an outcome evaluation:

- ➤ Limit the outcomes to be examined to only the most important ones for your program.
- ➤ Choose outcomes that can be measured using less costly data collection methods.[6]
- ➤ Choose indicators for which data already exist.

Impact evaluations require an even higher level of financial and technical resources.

Impact evaluations should only be undertaken when there is a compelling reason for doing so, such as to demonstrate the efficacy of a program strategy in a particular target population, or to meet government or donor requirements.

[6] Data collection methods are discussed in Chapter 7.

Whatever resources you have, be creative in using them.

There are many ways to collect data. Programs often collect too much data, either collecting data about too many issues, collecting data that does not relate to their objectives or activities, or using different methods to collect the same data from the same target population. Spending a lot of resources on data collection does not guarantee that you will end up with results that help you better understand your program and participants.

Managing an M&E effort requires planning and creativity. Think carefully about the types of information you need to collect. Find ways to collect data that relate to the outcomes you hope to achieve, the meaning of your program for participants and the factors that influence why your program is succeeding.

Budgeting for an M&E effort is an important part of planning.

The worksheet on the following page will help you think about how to calculate the costs of each category in an M&E budget and can be used as a reference when preparing detailed estimates for each budget item. However, some decisions—such as what indicators and data collection methods will be used, and the frequency and timing of data collection—will be based on material discussed in Chapters 4–8 and should also be considered before you finalize your M&E budget.

Who Should Be Involved in M&E?

M&E efforts should involve many stakeholders, as many people in the community have an interest in M&E. Stakeholders may include program staff, youth, school administrators and teachers, parents, community leaders, local government officials, service providers and donors. They may be active or want to be involved in some or all phases of an evaluation: planning and design; collecting and analyzing data; identifying the key findings, conclusions and recommendations of an evaluation; disseminating the results; and, finally, planning how evaluation results can be used to improve a program.

It is important to involve staff and stakeholders such as community members and youth in the discussion of how M&E information will be used.

Stakeholder involvement can make M&E efforts more relevant and effective.

Participatory evaluation facilitates the identification of local needs and priorities, and places evaluation issues in the context of people's lives. Involving stakeholders can help you achieve the following M&E goals:

- Develop consensus about the key issues to be addressed in an evaluation.
- Identify what information stakeholders need about the program.
- Ensure that program staff understand the need for evaluation, their role in its implementation and how the results will be used to improve the program.
- Avoid intrusive or inappropriate evaluation methods.
- Create open lines of communication among stakeholders for later dissemination and discussion of evaluation results.

Worksheet 3.2
Preparing an M&E Budget

Item	Amount of Funds Needed	Source of Financial Support	Source of In-Kind Support
Salaries: For the personnel needed for technical assistance, data collection, data entry and analysis (staff, interviewers, supervisors, drivers, etc.)			
Per diem: Daily costs for lodging and food Travel: Bus or taxi fares, gasoline, vehicle rental and maintenance			
Printing: Survey questionnaires, interview guides, final reports, etc.			
Equipment: E.g., bicycles and computers (and including maintenance)			
Communication: Telephone, fax, computer, radio, postage, etc.			
Supplies: Paper, computer diskettes, pencils, portfolios, etc.			
Dissemination activities: Seminar or conference costs, refreshments, materials, portfolios, presentation supplies, etc.			
TOTAL			

Stakeholders can also help increase the knowledge of external evaluators about the program context and develop opportunities for continued contact between those conducting the evaluation and those affected by it.[7]

Participatory evaluation is one way to involve the most important stakeholders—youth.

Young people targeted by the program are its most important stakeholders. However, some adult program managers and staff may find it difficult to work with youth on a regular basis, given the many differences that can exist between the generations in terms of attitudes, behaviors and beliefs.

Participatory evaluation is a set of techniques that emphasize community involvement in gathering knowledge and help place issues of concern in the context of people's lives. This experiential knowledge aids in directing appropriate responses and defines the array of services offered. Participation generally takes place throughout all phases of the evaluation: planning and design, gathering and analyzing data, disseminating results and preparing an action plan to improve program performance.[8] Program planners in the United States have found some effective strategies for working with youth that have application across many social settings, and which are presented in the box at right.

Tips for Involving Youth in Participatory M&E[9]

Tip	Examples
Integrate young people into program efforts and M&E planning.	Schedule meetings in accessible locations. Maintain communication and convey needed information. Encourage full participation and voting rights.
Be open and nonjudg-mental about young people' s insights and suggestions.	Guard against dismissing or reacting negatively to young people' s suggestions. Make time for them to feel comfortable and participate fully. Solicit their ideas and opinions.
Take advantage of the expertise young people offer.	Encourage youth to share their knowledge and perspectives about positive or negative program effects.
Be honest about expectations for the program, young people' s contributions and benefits of youth participation.	Do not claim that the program can solve all problems. Be realistic about what you can tackle.
Offer support for young people.	Provide mentoring, financial assistance, transportation, training, supervision and information.
Make work interactive and fun.	Be creative and allow youth to be creative. Design programs that are informative, fun and fulfilling.
Help build young people' s skills so they can become more involved.	Provide information and build skills that increase youth' s confidence. Allow them to practice ways to communicate with different audiences.

Involving stakeholders and youth can raise problems.

Disadvantages to involving stakeholders, especially those from other organizations, include the following:

- ➤ It may be difficult to be objective in selecting representative young people and organizations to participate in the evaluation.
- ➤ Stakeholders may not know much about how a program works.
- ➤ Organizations may hold competing perceptions and concerns that are difficult to resolve or prioritize.

[7] Lawrence, 1989.

[8] USAID CDIE, 1996.

[9] Adapted from Clark, Haughton-Denniston, Flinn, et al., 1993, cited in Brindis and Davis, 1998b: Volume 4, p. 49.

- ➤ The ability of an evaluation to be independent may be compromised by including diverse organizations.
- ➤ More participants may require a greater allocation of your staff time and resources.[10]

WHO SHOULD CARRY OUT THE EVALUATION?

Evaluations can be done by your own staff, by those outside your program or by a combination of the two. When deciding who will carry out an evaluation, you should consider several issues. First, what is the most appropriate structure for the evaluation team? Second, what is feasible? What are you able to afford, given your budget? You may find that it is simply too cumbersome or inefficient to involve all stakeholders in every M&E activity.

Funding agencies sometimes require that evaluations be carried out, at least in part, by outside evaluators.

Using staff to carry out evaluation has advantages.

In-house staff members are familiar with the program and can be trained quickly. They also may be aware of particular program strengths or weaknesses that require attention. Finally, the results of the evaluation will be most useful to program staff, who are positioned to modify and improve the program accordingly. Using staff may also be more financially feasible, as outside evaluators are often more expensive. Also, for financial or logistical reasons, outside evaluators may only be available for a limited time.

Using outside evaluators is more appropriate in some situations.

Funding agencies sometimes require that evaluations be carried out, at least in part, by outside evaluators. Since they have less stake in the outcome of the evaluation, outside evaluators are perceived to be more objective in drawing conclusions and tend to have more credibility. However, while maintaining objectivity, outside evaluators must be sensitive to program goals and the local context within which the program is implemented. Rather than posing a threat, evaluators should be considered in light of their role as part of the support system for the program.

When staff resources are limited, using outside evaluators may be more feasible.

Whether to use in-house staff or outside evaluators also depends on the available time and expertise of your program staff, as evaluations can be very demanding. You will need to assess the experience and skills of your staff in conducting M&E, and how much time they will have to spend on these efforts. You will also need to consider which staff must be involved, how vacations and holidays may affect their availability, and whether you need any outside help. Ideally, youth and other significant stakeholders should participate to the extent possible. In some cases, evaluation may be coupled with technical assistance as part of a broader approach to enhance the effectiveness of the program and to train in-house staff.

[10] Lawrence, 1989.

WHERE SHOULD M&E TAKE PLACE?

If your program has only one or two sites or covers a small geographic area, you can more easily conduct monitoring and evaluation efforts for the entire area or set of sites. However, if your program covers a larger area or multiple sites, you may need to narrow the geographic scope of the effort. How you select the sites or areas to be included in your M&E effort will depend on your information needs and financial and human resources.

Make an effort to monitor each program site.

As monitoring is essential for effective program management, you should try to include all program sites in the collection of basic information—such as whether planned activities have been completed and the number and sex of clients that have been served by your program. This will provide a picture of how program implementation is progressing, as well as allow you to compare how sites perform in relation to one another. If some of your program sites have greater capacity to collect data than others, you might consider having them gather additional monitoring data that will be helpful in answering other questions about program implementation.

If it is not possible to collect the same monitoring data from all sites, you probably should not implement a program there unless you are absolutely sure that the strategy will work without monitoring. For many strategies—for example, peer education— monitoring is essential to ensure that the program is being implemented as planned. If you determine that monitoring is not needed for a program to work well, you can choose to monitor only parts that you think are "representative" of the sites in your program. How to choose a representative sample of sites is discussed in Chapter 6.

You probably will have to limit data collection for evaluation.

For example, it is rarely possible in process evaluations to evaluate every service contact or obtain feedback about the program from every participant. In larger programs, you might also have to limit process evaluations to only a sample of your program sites.

The advantage of using in-house staff members is that they are familiar with the program and can be trained quickly.

Most outcome and impact evaluations also require that some restrictions be applied about where data will be collected. This is especially true when program objectives pertain to outcomes measured for the general population of youth. Here, it will usually be necessary to collect data from youth in only a sub-set or sample of the geographic areas covered by a program.

Choosing sites for conducting evaluations requires careful consideration.

Sometimes a program does not have a clearly defined geographical area of influence. If the area of influence of a program is defined very broadly, such as an entire city or region, then it may be more difficult to measure changes in objectives even if the program performs optimally.

In choosing sites or geographic areas for conducting evaluations, ask yourself these key questions:

- ➤ What geographic area does the program reach?
- ➤ How many sites or geographic areas do I need to conduct a strong evaluation?
- ➤ Do these sites represent the characteristics of the youth target population and the program being implemented?
- ➤ How many observations do I need per site or geographic area?
- ➤ How should I go about choosing the sites or geographic areas?

Many of these issues are addressed in Chapters 5 and 6. The following practical considerations will also likely influence your decision about where to evaluate:

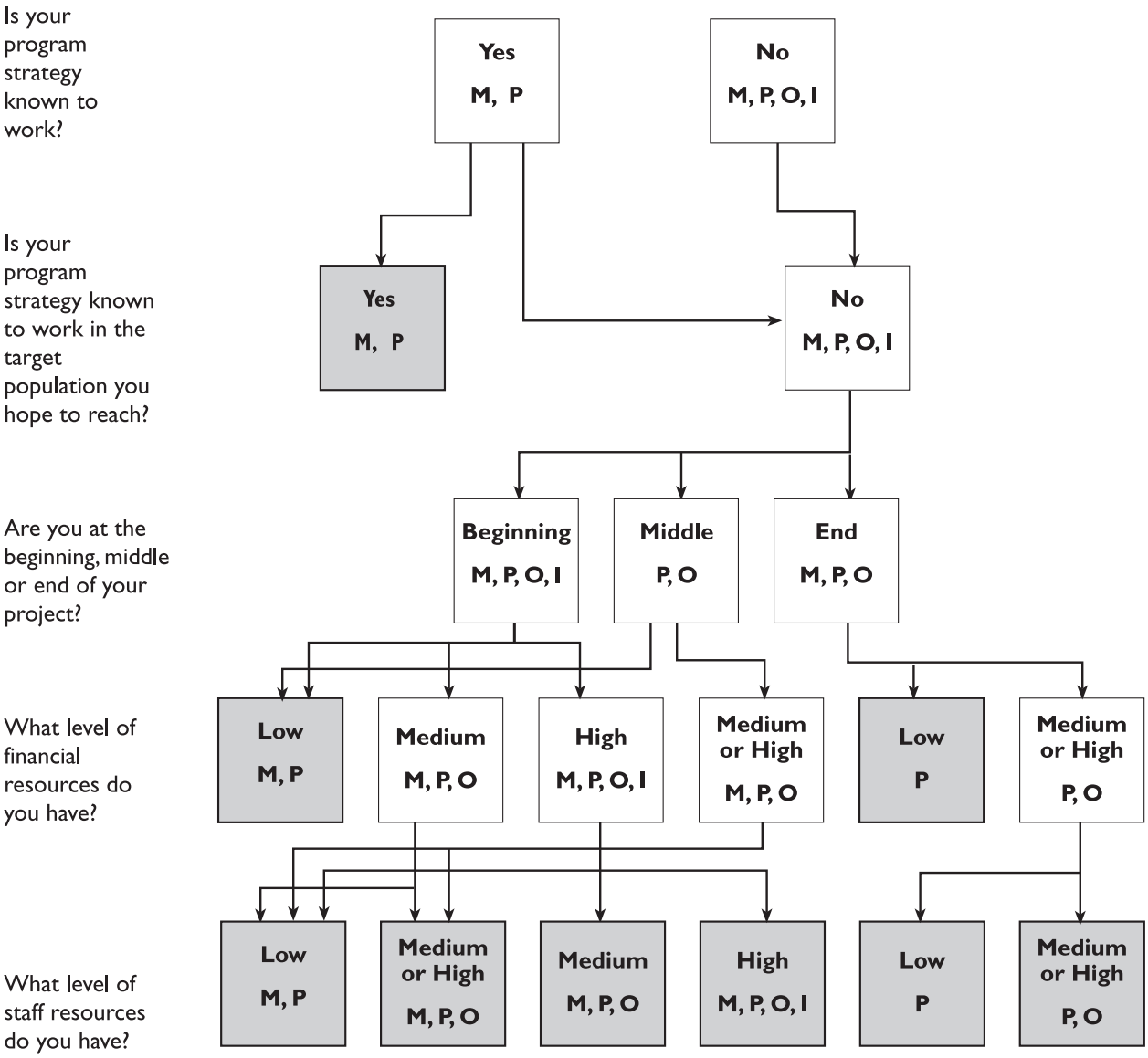
- ➤ Are existing data available in all sites?
- ➤ How easy is it to collect new data in each site?
- ➤ How will data collection affect the performance of regular program activities?
- ➤ Are there any other resources available to help collect and analyze the data (e.g., local universities or research groups)?

Determining the Type of M&E Effort You Undertake

The following checklist and flow chart can be used to help you determine the type of M&E activity you might undertake. Complete this checklist before using the flow chart:

- ❑ Are the goals of your program clear?
- ❑ Are your objectives related to your goals and intended outcomes?
- ❑ Are your objectives expressed in measurable terms?
- ❑ Are your activities defined?
- ❑ Do your activities relate to your program objectives?

Use this flow chart to determine the type of M&E effort you should undertake. When you get to a box in bold, this is the most appropriate type of M&E effort for your program.



M = Conduct Monitoring
P = Conduct Process Evaluation
O = Conduct Outcome Evaluation
I = Conduct Impact Evaluation

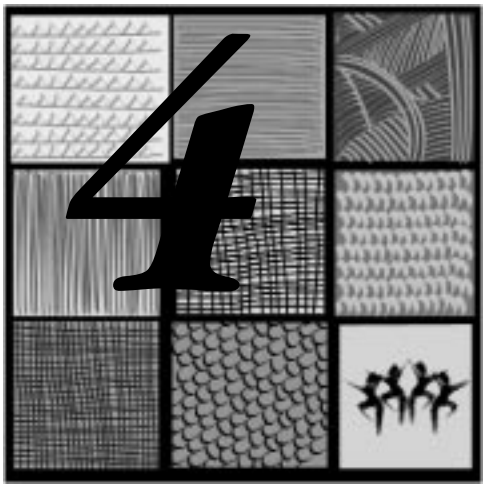
What Is Involved in Carrying Out Each Type of Evaluation? (How to Use the Rest of This Guide)

The table on the opposite page will help you determine how to use the rest of this Guide. Depending on the scope of your M&E effort, it guides you to relevant ARH indicators (Chapter 4 and Indicator Tables), data sources and data collection methods (Chapter 7) and corresponding instruments (Part II of this Guide). For outcome and impact evaluations, study designs are suggested in Chapter 5. The table refers to relevant sampling issues, which are explained in Chapter 6, and to types of analysis, described in Chapter 8.

How to Use the Rest of This Guide							
	Level of Resources Needed	**Selecting Indicators (Chapter 4)**	**Data Sources, Methods and Collection (Chapter 7)**	**Instruments and Data Collection Tools (Part II of this Guide)**	**Study Design (Chapter 5)**	**Sampling (Chapter 6)**	**Data Collection and Analysis (Chapter 7 & 8)**
Monitoring	Low	Indicator Table II: Program Systems Development and Functioning Indicators; Indicator Table III: Program Implementation Indicators	Service statistics; administrative reports and documents; event logs; other types of logs; surveys	Instrument 2: Tally Sheets; Instrument 3: Reporting Forms; Instrument 5: Composite Indices; Instrument 6: Inventory of Facilities and Services	N/A	Limited use	Comparing systems development and functioning and implementation indicators against targets; comparing indicators for different program sites; assessing trends of indicators over time
Process	Low to Moderate	Indicator Table I: Program Design Indicators; Indicator Table II: Program Systems Development and Functioning Indicators; Indicator Table III: Program Implementation Indicators	Administrative reports and documents; event logs; other types of logs; document review; site visits; direct observation; interviews with key informants (e.g., service providers, managers); unstructured feedback from clients; exit interviews with clients; mystery clients; focus groups or informal listening sessions	Instrument 1: Checklists; Instrument 2: Tally Sheets; Instrument 6: Inventory of Facilities and Services; Instrument 7: Observation Guide for Counseling and Clinical Procedures; Instrument 8: Interview Guide for Staff Providing RH Services; Instrument 9: Guide for Client Exit Interview; Instrument 10: Questionnaire for Debriefing Mystery Clients; Instrument 14: Assessing Coalition Effectiveness Worksheet	N/A	Choosing samples of program sites, participants, service providers, service transactions, etc. for the measurement of systems development/ functioning and implementation indicators	Comparing systems development and functioning and implementation indicators against targets or standards; comparing indicators for different program sites; assessing trends in indicators over time
Outcome	Moderate	Indicator Table IV: Program Intervention Outcome Indicators	Service statistics; surveys; population surveys; focus groups or informal listening sessions	Instrument 3: Reporting Forms; Instrument 6: Inventory of Facilities and Services; Instrument 11: Community Questionnaire; Instrument 12: Comprehensive Youth Survey; Instrument 13: Focus Group Discussion Guide for In-School Adolescents	N/A	Choosing samples of youth, program sites, communities, etc. for the measurement of outcome indicators	Comparing outcome indicators against targets; comparing indicators for different program sites; assessing trends in indicators
Impact	High	Indicator Table IV: Program Intervention Outcome Indicators	Service statistics; surveys; population surveys; focus groups or informal listening sessions	Instrument 3: Reporting Forms; Instrument 6: Inventory of Facilities and Services; Instrument 11: Community Questionnaire; Instrument 12: Comprehensive Youth Survey; Instrument 13: Focus Group Discussion Guide for In-School Adolescents	Chapter 5	Choosing samples of youth, program sites, communities, etc. in each experimental or comparison group for the measurement of outcome indicators	Comparing outcome indicators for "treatment" and "control" groups

PHOTO: Harvey Nelson

INDICATORS



CHAPTER AT A GLANCE

- ➤ Defines and explains indicators
- ➤ Provides examples of how to select and modify indicators to match your program objectives and activities

What Is an Indicator?

An *indicator* is a measurable statement of program objectives and activities. Once you have defined a program's objectives and activities, you can develop indicators—or measures—for each objective and activity. Some programs may have single indicators, and others have multiple indicators. Generally, it is preferable to have several indicators to capture the multiple dimensions of your program. However, you should carefully select a manageable number of indicators so that they accurately reflect your program objectives and activities and your evaluation priorities.

Continuing with the example presented in Chapter 3, the table below shows some indicators that can be used to measure the objectives and activities associated with delaying the age at sexual initiation through a peer education program.

Indicators can be expressed in different forms.
As you can see in the example above, indicators can be expressed in different ways. Numeric indicators are expressed as counts, percentages, ratios, proportions, rates or averages. The following indicators are counts:

- ➤ Number of radio advertisements aired
- ➤ Number of clients who seek peer counseling services

In evaluation terms, it is usually more informative to state indicators as percentages, ratios and proportions. These measures allow you to see what was achieved in relation to the *denominator*, or total possible number, while counts simply give you an idea of the number of events that took place, or the number of people reached, without indicating the total possible number. For example, you may count the number of youth who have delayed sexual initiation, but if you have a denominator, i.e., the total number of youth in a given geographic area, you will be able to calculate the proportion of youth in that area who delayed sexual initiation. This will allow you to measure the coverage of your program and the effects on behaviors at the population level.

Note

Later in Part I of this Guide, we provide a definition for each of these terms and give instructions for how to calculate different types of numerical indicators.

Determining Objectives, Indicators and Activities			
Objective	**Possible Outcome Indicators**	**Activities**	**Possible Program Indicators**
Population-level objectives: • Delay age of sexual initiation among youth ages 14– 19	• Average age of sexual initiation among youth ages 14– 19	• Promote availability of peer counseling services through radio ads	• Number of radio advertisements aired • Number of new radio advertisements aired
• Increase the percentage of youth ages 14– 19 who seek counseling services from peer educators to 25 percent	• Percentage of youth ages 14– 19 who seek counseling services from peer educators	• Implement " peer counseling corner" in five health clinics	• Number of clinics that have peer counseling corners • Number of days peer counseling corner is staffed per week • Number of clients who seek peer counseling services
		• Have peer educators give informational talks at schools twice a week	• Number of informational talks by peer educators in schools • Number of youth who attend informational talks • Quality of peer educators' presentations (based on criteria for curricula)
		• Have peer educators provide quality counseling services	• Proportion of clients who rate counseling as high-quality during exit interviews • Quality score of peer counselors (based on counseling criteria) given by observers
Program-level objectives: • Increase capacity of peer educators to provide counseling to youth	• Number of peer educators who are competent to provide counseling to youth	• Recruit peer educators from pool of adolescents who attend clinic	• Recruitment completed? (Yes/No)
		• Select 30 peer educators	• Number of peer educators selected
		• Develop training curricula	• Training curricula developed? (Yes/No) • Number of " key topics" training curricula covers as compared to checklist
		• Train peer educators to provide counseling	• Proportion of peer educators who demonstrate effective counseling skills during role plays

Non-numeric indicators are expressed in words. They are also referred to as *qualitative* or *categorical* indicators. These indicators usually denote the presence or absence of an event or criteria. The following are non-numeric indicators:

- Peer education recruitment completed? (Yes/No)
- Training curricula included topic on relationships and sexuality? (Yes/No)

Non-numeric indicators can also be used to summarize descriptions or assess quality or comprehensiveness. You can do this by creating an index of items that can each be assigned a number, which are then totaled to produce a score. In the table below, for example, each of the items in the right column would be assigned a point, and then those points would be totaled to determine the overall score of the presentation.

Using Non-numeric Indicators to Measure Quality

Indicator	Index and Quality Score
Peer educator's presentation is comprehensive	Observe presentation. Check each topic that is covered accurately. Give one point for each item checked, and total to determine quality score. ❑ Anatomy and reproduction ❑ Abstinence ❑ Contraception ❑ How to use a condom ❑ Making the decision to have sex ❑ How to say "no" to sex ❑ Resisting peer pressure to have sex ❑ Where to get counseling ❑ Where to get health services Total: _____

Like objectives, indicators should be specific.
The more specific your indicator, the more likely that you will accurately measure your objectives and activities. Indicators should specify the:

- characteristics of the target population you intend to reach, such as gender, age and residential, marital and schooling status;
- location of the target population, such as rural or urban youth, youth in a certain city or district, youth who participate in your program or youth who attend certain schools or clinics; and
- the time frame within which you intend to achieve your objectives.

An indicator should have the same scale as its corresponding program objective.
For example, if your objective is to delay the average age at sexual initiation among youth ages 14–19 who live in your district, then the indicator should measure "average age at sexual initiation among youth ages 14–19 who live in district *X*." If your indicator's scale is different from your objective's, your results will be misleading.

Types of Indicators

Once you have decided on the scope of your M&E effort, different indicators should be developed for each component of the program to be measured. For example, if you plan to conduct a process evaluation, you should develop indicators for design, systems development and functioning, or implementation. If you plan to conduct an impact evaluation, you should develop indicators for program implementation and outcomes.

In this Guide, we have categorized indicators based on what component of the program will be monitored and evaluated. Chapter 10 provides four Indicator Tables, each containing examples of indicators based on program aspects. You can use these tables to select and adapt indicators to match your program.

Design indicators are related to "key elements."
Youth programs should be designed based on "key elements" of quality. The international experience of youth programs; lessons from the field of maternal-child health, family planning and HIV/AIDS; and practitioner intuition and experience have produced a number of recommended key elements of youth program design. Some examples are:

- ➤ existence of clearly defined goals and objectives,
- ➤ involvement of local stakeholders in program planning, and
- ➤ assessment of needs and preferences of the target young adult audience for reproductive health services.

Systems development and functioning indicators are related to programmatic objectives and activities.
Programmatic objectives state results in terms of the organizational structure, management or operations of a program, and the corresponding activities involve the development and functioning of your systems. Systems development and functioning indicators measure whether an organization's or program's systems are operating and how effectively they have prepared program personnel for implementation. Examples of systems development and functioning indicators include:

- ➤ number of peer educators trained to provide youth counseling,
- ➤ existence of a clear organizational structure, and
- ➤ number of partnerships, networks or coalitions established to support the ARH program.

Implementation indicators are related to both programmatic and population objectives and activities.
Both programmatic and population objectives will be met by the implementation of program activities. Implementation indicators measure whether and how many planned activities have been conducted, and the quality of the implementation of those activities. Examples of implementation indicators include:

- ➤ number of youth who seek peer counseling services,
- ➤ number and type of involvement by stakeholders in the ARH program, and
- ➤ number and type of communication products developed for the target audience.

Outcome indicators are related to population objectives.
Population objectives state results in terms of the program participant and are measurable statements of the outcomes you hope to achieve in your target population. Outcome indicators measure the changes in outcome that your program's activities are trying to produce in your target population. Examples of outcome indicators include:

- ➤ average age at sexual initiation;
- ➤ percent of youth who say they would advocate healthy behaviors among their peers and friends;
- ➤ pregnancy rate among female youth during a specified time period; and
- ➤ incidence rate of STIs for young adults during a specified time period.

How Should Indicators Be Stated?

Precision and clarity about your indicators will produce meaningful results from your M&E effort.

Assess indicators in terms of their importance and ease in data collection.
Indicators are considered of high importance if one or more of the following applies:

- The indicator is a priority, given the purpose and scope of the evaluation.
- The indicator tests a new approach.
- Staff members want to know about the indicator.
- Youth have identified the indicator as important.
- A donor requires information that the indicator will measure.

If you determine that the data needed to calculate your indicators are not available, then new information will need to be collected. It is important to assess how easy or difficult the collection of these data would be. Factors to consider in determining ease of data collection are:

- sensitivity of topics (especially in terms of local norms and cultural context),
- staff resources and expertise,
- logistical requirements (e.g., transport, printing, vehicles),
- time,
- cost, and
- slang, vernacular and professional terms used to refer to subject.

State indicators in clear and precise language.
It is important to use clear and precise words and phrases to state your indicators. General indicators may be open to many interpretations and will hinder your ability to interpret M&E results. For example, a general indicator might be "Number of youth who seek peer counseling services." This indicator should be more precisely stated as "Number of youth ages 14–19 who reside in our district who seek counseling services from peer educators during a six-month period."

Outcome indicators measure the changes in outcome that your program's activities are trying to produce in your target population.

Avoid changing the wording of indicators after an M&E effort has begun.
Changing the wording of your indicators during program implementation may hinder your ability to interpret M&E results. For example, assume your indicator is "Number of youth ages 14–19 who reside in our district who seek counseling services from peer educators during a six-month period." If in the middle of your program you change this to count the number of youth ages 14–16 who seek counseling services, it may appear that the number of clients has gone down. Therefore, your results would suggest that fewer youth are utilizing your program, when in fact this may not be true.

If you have already begun your M&E effort and discover that your indicators are not specific enough, it is advisable to *add*

indicators rather than to change existing ones. For example, if you found that youth who seek counseling services are mostly between the ages of 12 and 15, you could add the indicator "Number of youth ages 12–13 who reside in our district who seek counseling services from peer educators during a six-month period." You would then continue to measure the original indicator for youth ages 14–19, in addition to the new indicator for youth ages 12–13.

Indicators should be consistent over time.
The indicators you use should be consistent for the duration of the monitoring and evaluation effort. If you drop, add or modify indicators during the program's implementation, then you may not be able to assess why changes are occurring in your target population. For example, consider the following indicator on STIs:

If you drop, add or modify indicators during the program's implementation, then you may not be able to assess why changes are occurring in your target population.

➤ Percent of young adults who report specific symptoms of STIs

Suppose that to measure this indicator, you initially developed a checklist of six symptoms that peer counselors use to record what their clients report. After six months, you review clinic records at four clinics in your catchment area, and find that youth whose symptoms are different from those on your checklist are being diagnosed with STIs. You then add another four symptoms to the checklist used by peer educators. This means that peer educators may begin to record youth who mention any of these four additional symptoms, whereas before these youth would not have been included. Therefore, if the percentage of youth reporting symptoms of STIs subsequently increases, you will not know if this change occurred because of a true increase in the prevalence of STIs, or simply because you added four more possible criteria to the checklist.

Carefully determine the time dimension of outcome indicators.
Most outcome indicators refer to medium- or long-term desired outcomes. For example, it may take several years to document changes in the pregnancy rate among female youth. What you define as medium- and long-term will vary according to the nature and complexity of the program's objectives and activities.[1] For example, some programs may define medium-term outcomes as those achieved within one year, and long-term outcomes as those achieved in five years.

You should make sure to establish a reasonable length of time to achieve desired outcomes. Youth programs are often under pressure to demonstrate outcomes and therefore try to measure changes in an unrealistic amount of time. Your results might then falsely indicate that you have not met your objectives. Once you determine the amount of time you think it will take to achieve your objectives, you can state the time dimension of your outcome indicators. You will then need to

[1] Many of the indicators included in the Indicator Tables at the end of Part I of this guide are medium-term (e.g., *No. of times YAs have had STIs in the past year*).

track your outcome indicators for a sufficient period of time to be able to observe changes.

Indicators should be valid and reliable. Indicators should be *valid*, which means that they accurately measure the concept or event they are supposed to measure. They should also be *reliable*, measuring the issue or event consistently every time. Assessing the validity and reliability of indicators helps to ensure that you minimize *error* in measurement.

Two steps can strengthen the validity of your indicators:

1. Develop indicators whose content adequately samples all possible meanings of a concept. For example, to measure the quality of interactions between youth and their parents, think about all the possible meanings of quality of interaction. You might determine that how often youth communicate with their parents, how long their conversations last, what topics they discuss and the young person's perception of the interaction all contribute to its quality. You therefore might develop a series of indicators that together measure the quality of interactions, such as:

   ➤ frequency of youth communication with parent over past week,

   ➤ average length of time of a parent-child communication,

   ➤ topics discussed by youth and their parents, and

   ➤ youth's perception of the quality of parent-child communication in the last week.

2. Develop indicators that explore the relationship between two measures of the same phenomenon. For example, in exploring a parent-child relationship you consider two related indicators:

   ➤ youth's perception of whether their parents understand them, and

   ➤ youth's perception of what types of problems they are able to discuss with their parents.

   By measuring both of these indicators, you would be able to assess the extent to which hypothesized relationships between related concepts can be verified. For example, you could measure whether all youth who say their parents understand them also say they are able to talk to them about a variety of their problems.

You can increase the reliability of indicators by reducing the chance that random, temporary conditions in a person, situation or set of measurement procedures occur:

➤ Check the consistency of an individual's responses by asking him or her similar questions more than once during a survey or interview. For example, a young man who reports having quality interactions with his parents but also says that he cannot talk to his parents when he has problems shows inconsistency in his answers. In data analysis, you could check to see how many youth gave similarly inconsistent answers. If many youth did, you would have identified an unreliable measurement of these indicators. If only a few youth did, you would have identified an error in the individual's understanding of these questions.

➤ Collect data at different times and check how consistent youth's answers are. For instance, you might ask the same series of questions about the quality of interactions with parents on surveys given every six months.

Worksheet 4.1
Preparing a List of Possible Indicators

1. Write your objectives in the table.
2. For each objective, write the activities you have planned to achieve the objective. Refer to the Logic Model you developed to ensure that activities that address all antecedent factors are included.
3. For each activity, note who will participate (for example, youth ages 8 to 12; boys; vulnerable populations) and where it will take place.
4. For each activity, refer to the Indicator Tables (Program Design, Program Systems Development and Functioning, Program Implementation and Program Intervention Outcome) to list all possible indicators, or develop your own indicators.

Objectives	Activities	Target Population	Location	Possible Indicators
Objective 1	Activity 1 Activity 2 Activity 3			Indicator 1 Indicator 2 Indicator 3 Indicator 4
Objective 2	Activity 1 Activity 2 Activity 3			Indicator 1 Indicator 2
Objective 3	Activity 1 Activity 2			Indicator 1 Indicator 2 Indicator 3
Objective 4	Activity 1 Activity 2			Indicator 1 Indicator 2 Indicator 3

- Assess the data you collect by looking for inconsistencies due to error in observation, coding or data entry processes. For example, check to see if youth interviewed by interviewers of different ages have significantly different answers. Also check to see whether answers to open-ended questions are coded correctly, for example, whether "happy" and "joyful" are coded as the same or a different response.

Rigorously testing validity and reliability may require outside assistance to perform statistical tests. Minimally, it is important that you consider these issues as you develop indicators.

Worksheet 4.2
Assessing Possible Indicators

1. List indicators from Worksheet 4.1 in the first column.
2. Clarify the scope of the program. Is it a large-scale effort to reach all members of the target population, or a smaller, more limited intervention that will reach only those who participate in specific services or activities?
3. For each indicator, write the possible sources of the data needed, such as survey or focus group.
4. For each source of data, circle whether data are available or will need to be collected.
5. Rate ease of data collection, based on availability, time and cost to collect.
6. Rate importance of indicator (high or low).
7. Determine priority based on ease of data collection and importance of indicator.

Possible Indicators (from Worksheet 4.1)	**Scope of Program** L = Large S = Small	**Are Data Available Now?** Y = Yes N = No	**Need to Collect New Data?** Y = Yes N = No	**Sources of Data**	**Ease of Data Collection** E = Easy F = Feasible D = Difficult	**Importance of Indicator** H = High L = Low	**Priority (1 is highest)** 1 = EH 2 = FH 3 = DH 4 = EL 5 = FL 6 = DL
Indicator 1	L S	Y N	Y N		E F D	H L	
Indicator 2	L S	Y N	Y N		E F D	H L	
Indicator 3	L S	Y N	Y N		E F D	H L	
Indicator 4	L S	Y N	Y N		E F D	H L	
Indicator 5	L S	Y N	Y N		E F D	H L	
Indicator 6	L S	Y N	Y N		E F D	H L	

PHOTO: JHU/CCP

EVALUATION DESIGNS TO ASSESS PROGRAM IMPACT

CHAPTER AT A GLANCE

- Offers guidance on an considerations around the need for impact evaluation
- Reviews study designs you can use to carry out an impact evaluation
- Outlines the technical requirements and resources needed for each type of evaluation
- Presents options for initiating evaluations after a program is underway

Why Should I Conduct an Impact Evaluation?

An impact evaluation will reveal the extent to which any observed changes in outcome indicators is due to your program activities. If your evaluation only measures changes in outcome indicators, your findings may not be fully credible for several reasons:

- **Other events or conditions may contribute to changes in outcome indicators:** Your program is only one of a number of factors that might affect the outcomes you are trying to influence. For example, changes in economic conditions or other social changes might influence how young people think about an acceptable age to begin having sex or use condoms. Other programs may be directed to the same target audience and conducted at the same time as your program. These type of external events could make the effects of your program appear to be larger or smaller than they really were. These factors are referred to as extraneous events.
- **Changes may take place within the individuals being studied over time:** Children's growth and development, or *maturation,* affects their attitudes and physical status, threatening the internal validity of an evaluation that aims to link changes in outcomes of such things as knowledge, attitudes or skills to health education or health promotion programs.
- **Program participants may have a predisposition to particular outcomes:** It is possible that your program attracts young people who are predisposed to the positive outcomes encouraged by the program activities. For example, your program might attract mostly youth with high education aspirations who might be less inclined toward risky behaviors. If so, simply measuring changes in outcome indicators would overstate the effectiveness of your program since many of your program participants would have realized

positive outcomes even without being exposed to your program. This problem is referred to as selection bias.

You should budget plenty of time before attempting to measure changes in outcomes, and ensure that your objectives clearly state the outcomes that you expect to produce.

Types of Study Designs for Impact Evaluations

The three major types of study designs for impact evaluations are:

- ➤ randomized experiments,
- ➤ quasi-experiments, and
- ➤ non-experimental designs.

(These different study designs are explained in detail later in this chapter.) In general, several factors differentiate one design from another:

- ➤ **Whether a "control" or "comparison" group is used:** A control or comparison group is a group of persons, facilities or communities similar to those who receive an intervention but who have not been exposed to the intervention. The purpose of a control or comparison group is to provide an estimate of what would have happened had you not implemented your program. A control group is randomly assigned, while a comparison group has similar characteristics but is not randomly assigned.
- ➤ **How participants are assigned to intervention and control groups:** In some evaluation studies, participants are assigned to intervention and control groups through random assignment. In others, control groups are selected—rather than randomly assigned—to match the characteristics of the intervention group, with the exception of their exposure to the intervention being evaluated.
- ➤ **The timing of data collection in relation to program implementation:** An evaluation may collect data before, during and/or after program implementation.
- ➤ **The complexity of statistical analysis required:** Some study designs require a more highly sophisticated statistical analysis.

There are several factors you should take into account when selecting an appropriate study design for measuring program impact:

- ➤ **Ethical issues:** Conduct a study only if it can be done ethically. If it compromises individual rights or denies the control group a chance at receiving a program or services they would have received if they did not participate in the study, change or abandon the study design. Note, however, that resource constraints often make it necessary to limit the target audience for a program, making the use of control and comparison groups more ethical. "Pilot" programs and phased-in programs provide opportunities to use experimental designs.

- **The importance of being able to demonstrate impact:** If community support and/or funding depends on demonstration of impact, you may be better served by using stronger study designs.
- **Validity:** *Validity* refers to the ability of a study design to measure the "true" impact of a program or intervention. The strongest study designs are those that are the least vulnerable to threats to validity. Two of the more significant threats to validity—extraneous events and selection biases—were mentioned at the beginning of this chapter. The "Summary of Study Designs for Impact Evaluation" box on page 87 summarizes the other major threats to validity in program impact evaluations, and provides guidance on choosing a study design that will be least vulnerable to threats to validity, given your particular circumstances and available resources.
- **Resource considerations:** The availability of funds, personnel, time and equipment will also influence your choice of study design. Your selection must be based on a realistic assessment of the resources available; what might be desirable with optimal resources may not be feasible in reality.

Randomized Experiments

Randomized experiments have the highest degree of validity among the evaluation designs. In evaluations using this study design, participants are assigned by chance (i.e., randomly) to a group that will receive an intervention (called the *intervention group*) or to a group that will not receive the intervention (called the *control group*). This is known as *random assignment* and is the best way to ensure that the members of the intervention and control groups are the same in terms of vital characteristics, such as age, education, family background, attitudes and any other factors that might influence the outcomes that the program seeks to influence.

The two main types of randomized experiments—pretest-posttest control group designs and posttest-only control group designs—are explained below.

Pretest-posttest control group design is the strongest design.
With the pretest-posttest control group design, you will:

- *randomly assign* persons, facilities or communities to experimental groups;
- take measurements both *before and after* the intervention; and
- measure impact as the *difference between changes in outcome indicators* for the intervention group and the control group.

The strongest of all designs, it is the "gold standard" among evaluation specialists. It does pose some challenges, however:

- It is not always feasible to randomly assign subjects to experimental groups.
- In order for randomized experiments to be effective, you will need to be able to maintain experimental conditions during the course of your program (for example, minimize the number of other programs introduced to either your intervention or control groups).

Ethical concerns
Is it ethical to withhold an intervention from some people in order to evaluate a program? When resources are sufficient to cover an entire population, this is indeed a difficult decision to make. However, in most communities, program resources are not sufficient to cover the entire population.

In such situations, *phased* or *delayed* intervention designs are often used. Phased program designs consist of giving an intervention to the group of youth who were studied as a control group in a subsequent phase of the program. This design goes a long way toward resolving ethical concerns, as well as permitting a strong program impact evaluation to be undertaken.

Two examples below and on page 75 present characteristics of large studies, perhaps larger than what may be feasible for your program. Regardless of size, however, these examples demonstrate the feasibility of evaluating reproductive health programs.

Evaluating a Program to Postpone Sexual Involvement: A Pretest-Posttest Randomized Experiment

What kind of program was being evaluated?
From 1992–94, evaluators studied the Postponing Sexual Involvement (PSI) program in California, USA, using a randomized pretest-posttest control group design.

What issues were being measured by outcome indicators?
The primary outcome indicators being measured were:

- beliefs about sexual activity,
- reasons to have sex or abstain,
- beliefs about sex and the media,
- personal communication,
- self-efficacy (e.g., confidence in one's ability to negotiate use of a condom when needed),
- behavioral intentions, and
- actual behaviors.

How were the experimental groups assigned?
A total of 10,600 youth were assigned to experimental groups as follows:

- In some sites, classrooms within sample schools were randomly assigned to youth-led intervention, adult-led intervention or no intervention (control) groups.
- In other sites, entire schools were randomly assigned to either intervention (adult-led) or no intervention (control) groups.
- Youth recruited from community-based agencies were randomly assigned to either intervention (adult-led) or no intervention (control) groups.

When were data collected?
Survey data were collected at three points: prior to program implementation, 3 months after program implementation and 17 months after program implementation.

How was impact measured?
Impact was measured by comparing changes in outcome indicators for youth in each of the three sets of experimental groups at 3 months and again at 17 months after program implementation.

What was the impact?
There was no evidence that the program influenced knowledge, attitudes or behaviors that the program attempted to change. Minor trends were observed at 3 months but were not sustained to 17 months.

Adapted from Kirby et al., 1997.

HIV Prevention in Tanzania:
A Pretest-Posttest Randomized Experiment

What kind of program was being evaluated?
From 1991–94, evaluators studied an HIV reduction program in Mwanza, Tanzania, that aimed to improve screening and treatment of sexually transmitted infections (STIs). This evaluation, which included a total of 12,537 persons, used a randomized pretest-posttest control group design.

What issues were being measured by outcome indicators?
- HIV sero-conversion rates
- STI prevalence
- Prevalence of selected risky sexual behaviors

How were the experimental groups assigned?
The following steps were taken for random assignment of the study population:
- Six pairs of matched communities (in terms of location and socioeconomic factors) were chosen for the study (12 communities in total).
- In each of these six pairs, one community was randomly assigned to the intervention group and the other to the control group.
- In each community, a random sample of 1,000 households was chosen, and all persons age 15 and above were included in the sample.

When were data collected?
Biological and survey data were collected at two points: prior to program implementation and again 24 months later.

How was impact measured?
Impact was measured by comparing changes in outcome indicators for persons residing in "treatment" villages versus those in control villages over the 24-month study period.

What was the impact?
HIV infection rates over the 24-month period were significantly lower in villages that received STI screening and treatment, as opposed to villages with no intervention.

Adapted from Grosskurth et al., 1995.

Posttest-only control group design is somewhat weaker.
With the posttest-only control group design, you will:

- ➤ *randomly assign* persons, facilities or communities to experimental groups (same as pretest-posttest control group design);
- ➤ take measurements *only after* the intervention; and
- ➤ measure impact as the *difference between outcome indicators* for the intervention group and the control group, at some point after program implementation.

The key assumption in this design is that the experimental groups are equivalent in terms of all factors other than exposure to the intervention.

Posttest-only control group design is weaker than pretest-posttest control group design, because it doesn't take measurements before a program intervention. Posttest-only control group designs do not allow you to know the status of each group prior to program interventions. While you may be able to show some differences between the participants and the control group after program implementation, having measures of their status prior to the intervention enables you to make stronger conclusions about the program effects.

Checklist: Are We Able to Use a Random Assignment Study Design?

- ❑ The intervention we want to evaluate has not started yet.
- ❑ Staff or other professionals are skilled in conducting random assignment.
- ❑ There is available both a group of people who could be randomly assigned to receive the intervention and another group to participate as a control group.
- ❑ The financial resources and technical ability to collect data before and after the intervention begins are available.
- ❑ There are no ethical issues in withholding the program from the control group during the study.

Quasi-Experiments

Quasi-experiments use similar experimental groups, selected through non-random methods. If it is not feasible to randomly assign experimental groups, you can still take into account many of the external factors affecting your control and intervention groups by using a quasi-experimental design. This can be done by choosing a control group that is as similar as possible to the intervention group, often by *matching* on characteristics that are considered to be important antecedents of the outcomes sought by a program. For example, if you are conducting and evaluating a school-based program, you might choose control schools that are similar to intervention schools but not part of your program. Some of the characteristics you may consider when choosing control schools are:

- ➤ geographic location,
- ➤ grades taught,
- ➤ socioeconomic composition of the student body, and
- ➤ level of teacher training and experience.

Your control schools should then match the intervention schools on each of these characteristics.

Because matching is generally viewed as an imperfect substitute for random assignment, statistical analysis is usually used in quasi-experiments to take into account or "control for" differences in factors that could not be taken into account through matching. Usually, only a few of the factors on which comparison groups might differ can be taken into account by matching.

The three most commonly used types of quasi-experiments are:

- ➤ non-equivalent control group pretest-posttest design,
- ➤ non-equivalent control group posttest-only design, and
- ➤ generic control design.

Non-equivalent control group pretest-posttest design is a widely used type of quasi-experiment.
In fact, this may be the most widely used design in impact evaluations of social programs. It is similar to the randomized pretest-posttest control group design, except that the comparison group is created by matching rather than random assignment, which is why it is called "non-equivalent." The group members are not randomly assigned and do not have an equal chance of being assigned to the intervention or control group. With non-equivalent control group pretest-posttest design, you will:

- ➤ create experimental groups by *matching* particular characteristics that are considered to be important antecedents of the outcomes sought by the program;
- ➤ take measurements both *before and after* the intervention; and
- ➤ measure impact as the *difference between changes in outcome indicators* for the intervention group and the control group.

The major risk of this design is that the experimental groups might differ from one another on factors that cannot be accounted for through matching or statistical analysis, for example, factors that are difficult to measure (e.g., what proportion had experienced sexual abuse) or factors that you simply failed to measure.

Non-equivalent control group posttest-only design (static group comparison) is a somewhat weaker type of quasi-experiment.

This design is similar to the non-equivalent control group pretest-posttest design, except that a pretest is not included. With the non-equivalent control group posttest-only design, you will:

- ➤ create experimental groups by *matching* particular characteristics that are considered to be important antecedents of the outcomes sought by the program;
- ➤ take measurements *only after* the intervention; and
- ➤ measure impact as the *difference between outcome indicators* for the intervention and control groups, at some point after program implementation.

This design is often used when baseline data were not collected. It is a fairly weak design, however, as there is a high risk that the experimental groups might differ on factors that cannot be accounted for through matching or statistical analysis. When this design is used, more sophisticated statistical techniques will be needed in order to produce defensible estimates of program impact. However, even when sophisticated statistical techniques are used, there is still the danger that differences between experimental groups that were not measured during data collection will remain. For this reason, this design often produces impact evaluation results that are viewed with some skepticism.

Evaluating a School-Linked Health Facility in Brazil: A Non-Equivalent Control Group Pretest-Posttest Quasi-Experiment

What kind of program was being evaluated?

In 1995, in Salvador, in the State of Bahia, Brazil, the Secretariats of Health and Education began pilot-testing a school-linked health facility program intended to reduce pregnancy and STI transmission rates among young adults. The program featured family life education in schools, improvements in the "youth-friendliness" of services at public health clinics and formal linkages between schools and health facilities.

What issues were being measured by outcome indicators?

- Knowledge and attitudes
- Sexual behavior
- Youth development
- Environmental risk and protective factors

How were the experimental groups assigned?

A pretest-posttest quasi-experiment design was used to measure the impact of the program. Six program schools in metro Salvador were randomly chosen for the intervention group, along with six matched control schools. Schools were matched in terms of location, grade level and socioeconomic level of students.

When were data collected, and from where?

- Data were gathered in treatment and control schools prior to program implementation and then three years later.
- A "monitorng" survey was conducted in program schools at the end of the first year of the program. Students who changed schools during this period were tracked to their new schools.
- Health facility surveys—including observations of service transactions and exit interviews with young adult clients—were also undertaken at mid-program in order to assess the "youth-friendliness" of services.

What was the impact?

Impact will be measured to determine changes in responsible sexual behavior, negotiating skills and use of reproductive health services.

Adapted from FOCUS on Young Adults, final report forthcoming.

Evaluation of a Family Life Education Program in Peruvian Schools: A Non-Equivalent Control Group Posttest-Only Quasi-Experiment

What kind of program was being evaluated?

In 1995, the Peruvian Ministry of Education initiated a national family life education program in secondary schools. Because of limited resources, the program was to be phased in over a 5- to 10-year period. For a variety of reasons, it was not possible to undertake a baseline survey prior to program implementation.

In 1998, the Ministry decided to evaluate the program. Because no baseline survey had been conducted, a posttest-only control group design was chosen for the evaluation.

What issues were being measured by outcome indicators?

- Knowledge and attitudes
- Risk-taking and health-seeking behaviors

How were the experimental groups assigned?

A sample of 22 schools that had implemented the program was randomly chosen in Lima and in selected cities in the interior of the country. One matched "control" school was chosen for each pair of "program" schools, resulting in a total sample of 33 schools. The criteria used to match program and control schools included geographic location, size, age of school and socioeconomic level of the student body.

What data collection methods were used?

Data were gathered from students, teachers and administrators in each sample school, using self-administered questionnaires. Focus group interviews were also conducted with students and teachers.

What was the impact?

No systematic differences in knowledge, attitudes or behaviors were observed between treatment and control schools.

Adapted from FOCUS on Young Adults, final report forthcoming.

Generic control designs, the third type of quasi-experiment, compare changes in outcome indicators to the status of the general population.

Generic control designs can be used to assess whether changes or trends in outcome indicators for young adults exposed to your program differ from those in the general population of young adults. For example, imagine that you want to compare program data and national survey data regarding condom use among young adults. Using a generic control design, you might find that the proportion of your program target audience who used condoms in their last sexual encounter had increased significantly over the course of your program—but was still lower than the national level recorded among young adults.

You can use the generic control design only if there are data available regarding trends in *relevant* outcome indicators for the general population of young adults. You may be able to find this kind of data in large-scale surveys (e.g., Demographic and Health Surveys, Adolescent Reproductive Health Surveys) that are conducted on a periodic basis (e.g., every five years or so) in many countries. These surveys usually measure a number of outcome indicators relevant to adolescent reproductive health programs, such as age at sexual debut, frequency of sexual intercourse, number of partners, knowledge of risk factors for HIV/AIDS and condom use.

If you want to use the generic control design to measure the impact of your program, you must be sure of several things:

- ➤ The population for which control data are available must be similar in characteristics and composition to the target population for your program. This is the primary challenge of using generic control designs. For example, if your program targets high-risk young adults, it may not be appropriate to use the general population as a comparison group.
- ➤ The outcome indicators measured for the generic control population must be relevant to your program.
- ➤ Survey data must be collected from both groups at (or close to) the same time.

Non-Experimental Designs

Non-experimental designs do not use control or comparison groups. For this reason, these designs are a generally weaker means of measuring program impact than experimental designs. Non-experimental designs are used when:

- ➤ you have not made provisions for a control or comparison group as part of the evaluation plan,
- ➤ a program or intervention is expected to "reach" the entire target population of youth; this type of program is often referred to as a *full-coverage program* (e.g., intensive mass media campaigns and national family life education), or
- ➤ programs in settings where a high proportion of young adults remain in school to the secondary level.

The following sections describe three types of non-experimental design:

- ➤ time-series design,
- ➤ pretest-posttest design, and
- ➤ posttest-only design.[1]

Time-series design uses pre-intervention trends as a comparison.
The time-series design is the strongest of the non-experimental designs. With this design, you will:

- ➤ take *several measurements* of outcome indicators for a program's target population both *before and after* the intervention;
- ➤ use the indicators prior to implementation to project what would have happened in the absence of a program (assuming that this trend

Checklist:
Should We Use a Quasi-Experimental Study Design?

- ❑ The conditions do not exist to use a randomized assignment study design.

For All Types of Quasi-Experiments:

- ❑ We have identified the characteristics of youth that are important to match intervention and comparison groups.
- ❑ A group of youth exists who appropriately match the characteristics of the intervention group.
- ❑ Staff or other professionals are skilled in conducting matching.
- ❑ There are no ethical issues in withholding the program from the comparison group during the study.

For Pretest-Posttest Design:

- ❑ The intervention has not started yet.
- ❑ We have the financial resources and technical ability to collect data before and after the intervention begins.

For Posttest-Only Design:

- ❑ For intervention has already started, or a collection of baseline data is not feasible for some other reason (such as lack of resources).
- ❑ Staff or other professionals who can handle sophisticated statistical analysis are available.

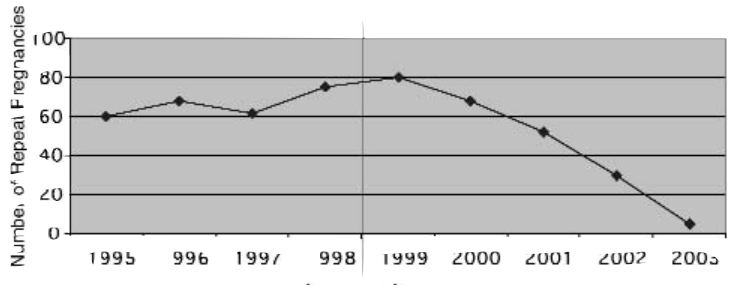
For Generic Control Design:

- ❑ Data exist about youth who have similar characteristics to those participating in the intervention group being studied.
- ❑ The available data measure the type of outcome indicators the intervention is trying to influence.
- ❑ Data was or will be collected from the generic control group at about the same time that data will be collected from the intervention group.

[1] These are classified by some as quasi-experimental designs, but because they do not involve comparison groups, we classify them here as non-experimental designs.

Illustrative Time-Series Analysis

In this example, the number of repeat pregnancies in the target population of all youth ages 15–19 in a specific geographical area was measured for a four-year period before the intervention began in 1999. Subsequent data collection showed that a year after the intervention began, the number of repeat pregnancies started to decrease. This downward trend can be associated with the intervention.



would have continued if a program had not been implemented);

- ➤ compare the trend in indicators prior to the intervention to the post-intervention trends; and
- ➤ detect program impact as a change in trends after an intervention was implemented (see the box below).

The time-series design has two major drawbacks. First, in order to use this design, you must have completed a sufficient number of observations prior to program implementation (at least three, and preferably more) in order to establish the trend that might have occurred in

Peer Education in Paraguay:
A Non-Experimental Pretest-Posttest Design

What kind of program was being evaluated?
From 1997–99, a media-based peer education program was implemented in the capital city (Asuncion) and in the central region of Paraguay. The objective of the program was to reduce pregnancy and STI transmission rates among young adults. The program featured a radio talk show, artistic performances and dissemination of relevant messages through other media channels. Peer educators were the performers/presenters in all activities.

Because the program was a full-coverage program (that is, it was expected to "reach" all young adults in Asuncion and in the central region), an experimental study design was not possible. Therefore, a non-experimental design was used.

What issues were being measured by outcome indicators?

- Knowledge and awareness of RH issues among youth
- How the media covers issues related to ARH

When were data collected, and through what methods?
Changes in outcome indicators were measured through surveys conducted among in-school and out-of-school youth prior to and two years after program implementation.

How were the data analyzed?
To assess whether observed changes in outcomes could plausibly be attributed to the program, behavioral indicators in the follow-up survey were related to the subjects' degree of exposure to program activities, events and messages. Sophisticated statistical methods were used to control for the fact that young people with higher levels of exposure to the program may have differed on significant characteristics from those with lower levels of exposure.

What was the impact?
Significant changes in selected knowledge, attitude, belief and intention indicators were attributed to exposure to the program, for example, an increase in the perception that girls who protect themselves are responsible. An increase in the proportion of youth reporting condom use at first sex was observed, but the change was not attributable to the program.

Adapted from FOCUS on Young Adults, final report forthcoming.

the absence of the program. (Because such pre-intervention data are rarely available, the time-series design is not widely used.) Second, it is difficult to account for the effects of adolescent maturation and other extraneous factors that may influence the outcomes during the period that your program was being implemented.

Pretest-posttest design is a more common type of non-experimental design. The pretest-posttest design is commonly used when you do not have a control group and you do not have enough pre-intervention data to use a time-series design. In this design, you will:

- ➤ measure outcome indicators for a program's target population both *before and after implementation*; and
- ➤ test any observed changes for statistical significance (that is, to determine if the changes were larger than what might be expected through random measurement error).

While this design allows you to document changes in outcome indicators in the target audience for your program, it is difficult to know to what extent these changes are actually due to your program. Extraneous factors could have either a positive or negative effect on your program's intended outcomes, thus hiding the program's true effect. If there are positive trends in outcome indicators, the most that you can say conclusively is that your program may have contributed to the positive changes. And in the case of negative impact evaluation findings, you will have little defense against the conclusion that your program was ineffective—even if the negative impact might have been caused by extraneous or outside factors.

One way to strengthen the findings from pretest-posttest designs is to look for *statistical associations* between study subjects' level of exposure to the program and changes in outcome indicators. For example, if you can show that young adults with high levels of exposure to a program had more favorable outcomes than did those with little or no exposure, then a stronger case can be made for the program having had an impact. The same is true if you can show that there were more favorable outcome indicators at program sites that had exhibited larger improvements in program implementation indicators than at sites with lesser improvements.

Note

Hiring external consultants to assist with your evaluation

Hiring external consultants to assist with your evaluation effort can be a good way to decrease demands on your staff, give credibility to your evaluation effort and compensate for skills that your in-house staff do not have. To ensure that your consultant provides the assistance your effort really needs, follow these tips:

- With staff and stakeholders, determine what tasks the consultant will undertake, and develop a clear scope of work. List the skills and strengths you would like the consultant to have.
- Interview several candidates and make sure that the person you hire has the skills you need and a shared vision of what the evaluation will accomplish.
- Develop a contract so that expectations are clear, spelling out what rights the consultant has in using the evaluation data collected.
- Before the consultant begins work, have him or her meet with staff and stakeholders to discuss their issues and concerns.
- Designate one staff person to act as a liaison between the program and the consultant.
- If the consultant is undertaking the entire evaluation, make sure that he or she debriefs staff and stakeholders on evaluation findings before publishing results.

Checklist:
Should We Use a Non-Experimental Study Design?

- ❑ The conditions do not exist to use a randomized assignment or quasi-experimental study design.

For All Types of Non-Experiments (only one reason need be checked):

- ❑ It is financially or logistically infeasible to collect data from a comparison group.
- ❑ The intervention is reaching the entire target population of youth.
- ❑ The intervention ethically cannot be withheld from the comparison group during the study.

For Time-Series Design:

- ❑ The intervention has not started yet.
- ❑ We have the financial resources and technical ability to collect data at least three times before the intervention begins.
- ❑ Staff or other professionals are skilled in projecting trends based on baseline data.

For Prettest-Posttest Design:

- ❑ It is infeasible to collect baseline data from the intervention group.
- ❑ Staff or other professionals are skilled in conducting more sophisticated statistical analysis.

Checklist:
Should We Use a Panel Study Design?

- ❑ We want to measure changes in some indicators more precisely.
- ❑ We have a good chance of being able to retain a high proportion of study participants for the duration of the study (e.g., study participants are not highly mobile or difficult to locate).
- ❑ We have sufficient resources to track study participants for future rounds of data collection.

These types of analyses can be rather complicated. If you use a non-experimental design of this type, you should probably seek help in carrying out the analysis from an external consultant, research firm or university faculty member.

Posttest-only design is the weakest of the non-experimental designs.
In the posttest-only design, data are collected for a program's target audience only after the intervention has been carried out. This is the weakest of all designs for measuring program impact, since it does not allow you to measure either:

- ➤ changes in outcome indicators over time, or
- ➤ differences between the target audience and other young adults not exposed to your program.

Because of the above weaknesses, many evaluation specialists would not even call this a study design. This design may be useful, however, for measuring aspects besides program impact, such as program coverage, the characteristics of youth served and not served by your program, and reasons why young adults might not use your program's services.

Panel Studies

Panel studies can be conducted as part of another study design. A *panel study* collects data from the same study subjects in each round of data collection. For example, if you are measuring changes in behavioral outcome indicators in a school-based program, you might collect data from the same "panel" of students at two or more points in time. Although panel designs often involve more than two rounds of data collection, any of the pretest-posttest designs described earlier could be considered a panel design if data were collected from the same "panel" of study subjects in baseline and follow-up surveys.

The alternative to using a panel design is to select a new sample of subjects in each data collection round. For example, instead of always collecting data from the same students in an evaluation of a school-based program, a new sample of students from the same sample of schools could be chosen each time.

There are both advantages and drawbacks to panel designs.
While panel designs allow you to measure changes in indicators much more precisely than in non-panel studies, they also can be difficult and costly. Following a panel of study subjects—especially higher risk youth—over a long observation period (for example, two to three years) requires patience and resources. If many participants drop out from the sample, bias can result.

In general, you should only use panel designs when:

- ➤ you have a good chance of being able to retain a high proportion of study participants for the intended duration of the study, and
- ➤ you have sufficient available resources to be able to track study participants for future rounds of data collection.

Tips for Selecting a Study Design

The points below offer some guidance on selecting a study design, in descending order from highest validity to lowest.

- Try to create treatment and control groups by assigning cases randomly from your target population (randomized experiments).
- If random assignment is not possible, try to find a comparison group that is as similar to the treatment group as possible. Check to see if the use of a generic control is possible (quasi-experiments).
- If neither a randomly assigned control group nor a similar comparison group is available, try to use a time-series design that can provide information on trends before and after program implementation (non-experimental designs).
- If sufficient pre-implementation data are not available to use a time-series design, try to obtain baseline (pretest) information that can be compared against post-program information (non-equivalent pretest-posttest design).
- If baseline information cannot be collected, be aware that your ability to produce strong conclusions about program impact will be limited. You will almost certainly have to use sophisticated statistical analytic techniques.

Adapted from Fisher et al., 1991.

Minimizing Threats to Evaluation Validity

While stronger study designs are generally more valid, there are potential threats to an evaluation's validity regardless of what study design you use. Following are the major threats to validity in impact evaluations and how you might minimize them.

How to Minimize Threats to Validity

Threat to Validity	How to Minimize Threat
History, or extraneous events that occur during the intervention period, can influence the outcome of the intervention. Changes in broad societal factors (such as economic, political or social conditions), extreme weather conditions, changes in public policies, or program-specific factors can affect outcomes. If, for example, a peer educator became pregnant and was no longer able to serve as an appropriate role model for your program, it may send a message to the target population that unprotected sexual intercourse is acceptable, compromising the ability of the program to produce its intended outcome.	• Select a study design with a control or comparison group. • Monitor extraneous events, such as mass media effects, that may influence your population in order to explain unexpected evaluation findings. • Undertake a process evaluation to understand how the program is working and take into account program-specific factors that may have affected outcomes during data analysis.
Selection bias is a difference between the people selected for the intervention and those in the control group. For example, youth living near one drop-in center may be wealthier, better educated and more likely to be involved in school or private club activities, and therefore may not be active in center activities. Meanwhile, youth living near another center may be out of school and have fewer opportunities for recreation, and may therefore use the center more frequently. This difference means that those exposed to the intervention are of a different socio-economic or educational background than those in the control group, and thus would respond to the intervention differently.	• Select a random sample of study participants from household listings. • Refer to other data sets about youth, compare the socio-demographic background of youth participating in the intervention to youth who are not, and include these findings in your evaluation report. • Ensure that you are matching the control group to the intervention group on as many characteristics as possible.
Testing prior to an intervention (either pilot-testing a survey instrument or giving a baseline pretest) is likely to affect the responses given in a posttest. People given a pretest (for example, youth interviewed before a media campaign stressing safe sex) may remember the questions and their answers when they respond to the posttest (after the media campaign has been run). Pre- and post-intervention differences in indicators might be due to familiarity with the questions, or youth having "learned" what responses are desirable, rather than to the effect of the intervention.	• Use a control or comparison group so that changes due to the intervention are more likely to be identified. • Pilot-test survey instruments with youth who will not participate in the study. • Change the order, but not the format or content, of items included on the posttest.

How to Minimize Threats to Validity	
Threat to Validity	**How to Minimize Threat**
Maturation occurs as time passes and may produce outcomes in participants that are unrelated to the effect of the program. For example, if participants become tired and bored during a lengthy training program, test scores may reflect fatigue and distraction rather than training effectiveness. Also, youth who are in the program may experience changes in attitudes and behaviors as they grow older, which are unrelated to the program.	• Make sure program implementers are aware of the need to minimize maturation effects as much as possible. • Use a control or comparison group to account for maturation. • Keep program interventions lively, diverse and as interesting as possible.
Dropout can affect evaluation results. If a program is implemented over a long period of time, participants may drop out, graduate, move or die, and so be lost to follow-up studies. If those that are lost differ substantially in significant characteristics from those who remain, the results of a post-intervention study may reflect those differences rather than the effect of program implementation.	• Compare characteristics of those who continue the program with those who drop out; if the two groups are not significantly different, you can show that dropout did not affect study results. • Ensure that you collect adequate contact information from youth studied over time. If possible, collect data from an institution with which youth are involved (e.g., a school) and ask respondents (1) where they think they'll be at the time of survey follow-up and (2) the contact information of one or two people who will always know where they are. • Budget enough to track youth who have moved or dropped out.
Instrumentation changes, such as modifying a survey questionnaire between the pretest and the posttest, can result in an effect independent of the project interventions. Deteriorating equipment, such as scales, rulers or calipers, may yield inaccurate readings. Some changes may also be attributable to improved skills on the part of the interviewer. Having more experience, interviewers may elicit more complete information from a respondent, resulting in changes that are separate from the effect of the intervention. Interviewers may also become careless as they become more familiar with the interview guide or checklist.	• Keep the exact wording of baseline questions. • If you decide to measure new indicators at follow-up, add entirely new questions to your survey instrument and analyze them with a posttest-only design. • Collect data on new indicators using a different data collection method (e.g., focus groups or client exit interviews). • Ensure that those collecting data are well trained before the first round of data collection. • Ensure that interviewers are unaware of participant and control group members' assignments, so they will not anticipate responses.
Changes in program implementation can also affect an evaluation's validity. For example, a program that is discontinued, scaled-up or changed in content may produce different outcomes from what was planned.	• Evaluate interventions that are well-established and are less likely to be changed or threatened because of external events, such as policy changes or loss of funding. • Make sure that those implementing the program understand the effects of changing the intervention, and have them document and justify any changes made. • Conduct a process evaluation to document how the program changed over the course of the implementation.

Choosing a Study Design for Ongoing Programs

Your options to initiate evaluation—and your ability to prove impact—become more limited as your program progresses. If you begin your evaluation in the middle or near the end of a program or program cycle, evidence of impact will probably not be as strong as it would if you had started your evaluation at the beginning. If you must choose a study design after your program has begun, consider the following:

Options for Choosing a Study Design

In the middle of a program:

If you collected baseline measurements of key outcome indicators:

- Continue to track these indicators in mid-term and final rounds of data collection. This will in effect be the same as a *non-experimental pretest-posttest design.* You should also check to see if repeated surveys are being conducted in your setting that you might use as a generic control population.

If you did not gather baseline measurements of key outcome indicators, you have two options:

- Collect data now, treat these as baseline data and conduct a follow-up round of data collection near the end of the program or program cycle. This will again be the same as a *non-experimental pretest-posttest design,* but will cover only part of the program period. You will miss the portion of your program's impact that occurred prior to the initial round of data collection, but if your program was implemented slowly this is less of a problem.
- Wait until the end of the project and use a *posttest-only design.*

As both of these options are quite weak designs, you should explore the possibility of either including a control group or using a generic control (if one is available in your setting). Having a control population of some kind would strengthen the evidence of your program's impact.

At the end of a program or program cycle:

If you collected baseline measurements of key outcome indicators:

- Measure these indicators again in a final round of data collection. This will be a *non-experimental pretest-posttest design.* You should also check to see if repeated surveys are being conducted in your setting that you might use as a generic control population.

If you did not gather baseline measurements of key outcome indicators, you have a number of different options—some more desirable than others:

- Use a *non-experimental posttest-only design* and measure indicators at the end of your program period. However, this is a very weak design; the following options are preferable.
- Include a control group (this would be a *non-equivalent control group posttest-only design*).
- Use a generic control (this would be a *generic control posttest-only design*).

Whichever option you choose, waiting until the end of a program or program cycle to collect data to measure impact will mean that:

- you will need to do fairly involved statistical analysis to be able produce credible impact evaluation results, and
- you run the risk of not being able to attribute changes in your target population to your program efforts.

Summary of Study Designs for Impact Evaluation

Class of Study Design	Defining Feature	Specific Study Design	Defining Features	Advantages	Weaknesses	Situations in Which This Design May Be Used
Randomized experiments	Study subjects are randomly assigned to treatment and control groups	*Pretest-post test control group*	• Random assignment • Measure indicators before and after • Impact shown as difference between changes in indicators for control and treatment groups	Strongest design—highest level of validity	• Random assignment is not always feasible • Must maintain experimental conditions throughout program	When you are starting a new program or program cycle
		Posttest-only control group	• Random assignment • Measure indicators only after • Impact shown as difference between indicators for control and treatment groups	Do not have to collect baseline data	• Weaker than pretest-posttest (no knowledge of pre-intervention status) • Must maintain experimental conditions throughout program • Less statistical power	When you are willing to trust that random assignment has created equivalent experimental groups (so that baseline data are not needed)

Summary of Study Designs for Impact Evaluation

Class of Study Design	Defining Feature	Specific Study Design	Defining Features	Advantages	Weaknesses	Situations in Which This Design May Be Used
Quasi-experiments	Control group is similar to treatment group, but not equivalent	*Non-equivalent control group pretest-posttest*	• Similar to "pretest-posttest control group," but experimental groups are chosen by matching rather than random assignment	Easier to choose groups through matching than through random assignment	• Must maintain experimental conditions throughout program • Risk that experimental groups may differ on factors that matching or statistical analysis cannot account for	• When you have collected baseline data • When random assignment is not feasible
		Non-equivalent control group posttest-only	• Similar to "posttest-only control group," but experimental groups are chosen by matching rather than random assignment	Do not have to collect baseline data	• Must maintain experimental conditions throughout program • High risk that experimental groups may differ on factors that matching or statistical analysis cannot account for • Less statistical power	• When baseline data were not collected • When random assignment is not feasible
		Generic control	Compare changes in program indicators to the status of the general population	No need to create control group	Data for relevant indicators and appropriate comparison population are rarely available—limited opportunities to use this design	• When indicators for generic control population are relevant to your program • When generic control population is similar to your program's target population

Summary of Study Designs for Impact Evaluation

Class of Study Design	Defining Feature	Specific Study Design	Defining Features	Advantages	Weaknesses	Situations in Which This Design May Be Used
Non-experimental designs	No control group	*Time-series*	• Measure indicators several times before and after intervention • Impact shown as a change in trends after intervention	Strongest of non-experimental designs, if have long enough time-series prior to intervention	Must have done at least three or more pre-intervention observations	• When you have not planned for a control group • When you are evaluating a full-coverage program
		Pretest-posttest	• Measure indicators for program's target population before and after • Test any observed changes for statistical significance	Can document changes in program outcome indicators in target population	• Difficult to attribute impact to your program • Requires sophisticated analysis to strengthen findings	• Same as above • When you don't have enough pre-intervention data for a time-series
		Posttest-only	Measure indicators for program's target population only after intervention	None for measuring impact—data may be used to inform new program design or as baseline for next program cycle	Weakest design—can measure neither changes in indicators over time nor differences with a control group	• Same as for time-series • When you have no pre-intervention data • When there is a need for post-intervention descriptive data

PHOTO: Harvey Nelson

SAMPLING

CHAPTER AT A GLANCE

- ➤ Describes types of sampling methods and ways to determine which one is appropriate for your program
- ➤ Focuses on one commonly used sampling method: cluster sampling
- ➤ Reviews how to determine and calculate the sample size you need for your data collection activities

What Is Sampling, and What Role Does It Play in Program Evaluation?

Program reports, service statistics and data collected on a routine basis can be used to measure some indicators within your evaluation. However, in certain cases you will need to collect data beyond existing records. Suppose, for example, that you want to measure changes in knowledge about AIDS among young people reached by the program. Collecting data from all youth involved in the program may not be possible, given the large number of people in question, as well as the limits of your program resources. It would be possible, however, to collect data from a *sample*, or sub-set, of youth. Sampling will enable you to collect a smaller amount of data that represent the whole group. This will save time, money and other resources, while not compromising on reliability of information.

Sampling is the process of systematically choosing a sub-set of the total population you are interested in surveying. With sampling, you produce findings that can be generalized to the target population of your program.

Sampling is useful in both process and outcome evaluations. For example, in conducting process evaluations, sampling enables you to:

- ➤ choose a sample of facilities to assess whether they have the necessary equipment and supplies to properly implement your program,
- ➤ choose a sample of health service transactions to assess whether service guidelines are being followed by service providers, and/or
- ➤ choose a sample of households or communities to assess the coverage of your program.

For outcome or impact evaluations, sampling enables you to:

- ➤ choose a sample of youth from your program or the general population to assess their levels of risk behaviors, and/or
- ➤ choose a sample of schools to assess whether students' life skills have improved as a result of your program.

Key Sampling Terms

Sample: A part of a whole selected to represent that whole.

Element: The person from whom you will collect data; an element could be a young person, a parent or a service provider.

Cluster: An aggregated group of elements from which you will collect data; a cluster could be a classroom, school, health facility or youth group.

Sampling frame: A comprehensive list of all relevant elements or clusters that is used to select a sample.

You may choose to seek outside help. Sampling, when done well, is a key aspect of an evaluation's credibility. However, sampling can be complicated, especially when evaluating large, complex programs.

After reading this chapter, you may want to seek technical assistance to make sure that the sampling scheme you develop is statistically sound. Experts at local universities and research institutes can help you develop an efficient sampling scheme for your program evaluation.

In general, information collected during a process evaluation will measure program-level objectives. Information collected during an outcome or impact evaluation will measure population-level objectives.

Types of Sampling Methods

Sampling methods fall into two broad categories: probability, and non-probability.
You will need to make a choice about which type to use. Probability sampling methods will strengthen the validity of your evaluation results. However, probability sampling methods can be time-consuming and costly and require a level of skill your program may not possess. Non-probability sampling methods are more flexible, less costly and less time-consuming, but give you weaker evaluation results.

Probability sampling methods are based on *probability theory*, a mathematical concept based on accepted statistical principles that refers to the ability to predict the statistical likelihood that a random event will occur. Probability sampling methods require that:

- every program element has a chance of being chosen (note: elements do not have to have an equal chance, only a chance),
- it is possible to calculate the probability of each element being selected, and
- random chance determines which elements are chosen.

Conversely, non-probability sampling methods are not based on probability theory. With these methods, samples are not chosen by random chance. These methods can include:

- obtaining a sample of subjects on the basis of opportunity,
- using your judgment to choose what you think is a representative sample of a larger population, or
- basing sample selection on referrals from other sample subjects.

The table below describes specific sampling methods that fall within the two categories.

Types of Sampling Methods	
Probability Sampling Methods	**Non-Probability Sampling Methods**
Simple random sampling: Elements are chosen at random so that each element has an equal chance of selection. For example, elements are chosen from a hat or, ideally, from a table of random numbers in a statistics textbook. They can also be generated by computer. **Systematic sampling:** The first element is chosen at random. Subsequent elements are chosen using a fixed interval (e.g., every tenth element) until you reach the desired sample size. **Stratified sampling:** The population to be sampled is divided into homogenous groups based on characteristics you consider important to the indicators being measured, such as youth that are sexually active. A simple random or systematic sample is then chosen from each group. **Cluster sampling:** First, a simple random sample of clusters is chosen from a sampling frame. Examples of clusters are schools, health facilities and youth clubs. Then, a simple random sample of individuals within each cluster is selected. **Multi-stage sampling:** This is like cluster sampling, but with several stages of sampling and sub-sampling. This method is usually used in large-scale population surveys.	**Convenience sampling**: A sample is drawn on the basis of opportunity. For example, the sample could include youth attending a school activity, service providers attending a conference or parents attending a school event. **Quota sampling:** A sampling frame is defined in advance of data collection and a sample is chosen from this list, but not at random. **Snowball sampling**: Data is collected from a small group of people with special characteristics, who are then asked to identify other people like them. Data is collected from these referrals, who are also asked to identify other people like them. This process continues until a target sample size has been reached, or until additional data collection yields no new information. This method is also known as *network* or *chain referral* sampling.

What Sampling Method Is Best?

To select the sampling method that is best for your needs, you should consider:

- ➤ the time and resources you have available,
- ➤ the size of your program,
- ➤ the indicators you plan to measure, and
- ➤ the data sources you plan to collect information from.

The box on the next page summarizes the major advantages and disadvantages of probability and non-probability sampling methods.

Probability sampling is preferred for most types of youth program evaluations.
If you work with a program that has a target population of more than 100 youth and are trying to measure changes in the indicators related to your program objectives, probability sampling methods will give you more objective and scientifically defensible evaluation findings. Non-probability sampling methods, while less costly and time-consuming, may leave you vulnerable to questions about whether your evaluation data are representative or unbiased. The primary reason some programs do not use probability sampling for their evaluation is because they lack the time, resources or expertise to conduct this type of sampling.

Non-probability sampling should be used in certain circumstances.
There are some circumstances in which probability sampling may not be possible or preferred:

- ➤ **The evaluation of smaller youth programs:** When there are small numbers of elements (fewer than 20) being chosen, the key statistical properties of probability sampling do not apply. If you plan to collect data from fewer than 20 individuals, a

Advantages and Disadvantages of Probability and Non-Probability Sampling Methods		
Type of Sampling Method	**Advantages**	**Disadvantages**
Probability Sampling	• Less prone to bias • Allows estimation of magnitude of sampling error, from which you can determine the statistical significance of changes/differences in indicators	• Requires that you have a list of all sample elements • More time-consuming • More costly • No advantage when small numbers of elements are to be chosen
Non-Probability Sampling	• More flexible • Less costly • Less time-consuming • Judgmentally representative samples may be preferred when small numbers of elements are to be chosen	• Greater risk of bias • May not be possible to generalize to program target population • Subjectivity can make it difficult to measure changes in indicators over time • No way to assess precision or reliability of data

non-probability quota sample that you consider representative in terms of important characteristics is just as effective as a probability sample.

➤ Some programs work with sub-groups of youth, such as street children or drug users. In these situations, it may be difficult or impossible for you to construct a *sampling frame*, or a list of all youth who belong to the target population. Without a sampling frame, you cannot conduct probability sampling. In these cases, snowball sampling is often the only feasible option.

➤ **When conducting focus group discussions:** Trying to assemble a randomly chosen group of respondents in a given location to conduct focus groups is difficult. Therefore, quota and convenience sampling are often used; of these, quota sampling is preferred because it is less prone to bias.

Cluster Sampling

Cluster sampling is the most widely used type of probability sampling and is useful for most types of ARH program evaluations. It involves two stages:

➤ First, identify groups that you want to collect data from, such as schools, health facilities or youth clubs. These groups are *clusters.* Next, make a list of all the clusters; this is your sampling frame. A sample of clusters is then selected from this list.

➤ Now develop a list of sample elements within your selected sample clusters, such as students, health facility clients or youth club members. Then select a sample of elements from each cluster. The result is the sample of individuals from whom data will be collected.

For example, suppose your organization has a peer education program in 40 schools. Each school would be considered a cluster. First you would select a sample of schools from among the 40 involved in the program. Then you would select a random sample of students within each school

included in the sample. This would be the sample you would collect data from.

Cluster sampling is the least costly type of probability sampling because it does not require you to develop a sampling frame of all elements before you select a sample. Developing sampling frames of individuals can be costly. For example, it might be difficult to develop a list of all clients who sought services at the 12 clinics you want to sample. It would be easier to list all of the clinics, each of which would be a cluster. After selecting a sample of clinics, you would only have to develop lists of clients of those clinics to be sampled. Once you have identified the target population and the indicators for your evaluation effort, you are ready to conduct cluster sampling.

Step 1: Define the clusters to be used.
A *cluster* is a clearly defined group of sampling elements from which you can select a smaller sub-sample. Some examples of clusters are:

- ➤ geographic areas with fixed boundaries (for household surveys—households and individual respondents are the sub-samples to be chosen in subsequent stages of selection);
- ➤ schools;
- ➤ health facilities; and
- ➤ youth clubs.

For example, clinics might serve as clusters for measuring process and outcome indicators. For measuring impact, the geographic areas prepared by a national statistics office (census enumeration areas, or CEAs) might be used as clusters.

The choice of what unit to use as a cluster will depend on what data are available. It is important to remember that with cluster sampling you can minimize the amount of effort needed to develop a sampling frame.

Sample Elements and Possible Clusters

Sample Element	Possible Clusters
Youth	• Geographic areas, communities, households • Schools • Classes, sections in school • Youth organizations, sports clubs • Health facilities • Workplaces • Youth named as a contact by a peer educator
Parents and extended family members	• Neighborhoods, villages, other geographic areas • Those whose young adult children have been selected for a sample (using same clusters as above)
Teachers	• Schools • Classes, sections in school
Community leaders	• Geographic areas, communities • Religious institutions • Commercial institutions • Government offices
Service providers	• Hospitals, clinics • Hotlines • Home visits
Health service events	• Hospitals, clinics • Pharmacies • Work shifts within a health facility
Peer educators	• Schools • Youth organizations • Workplaces
Commercial sex workers	• Brothels • Bars, clubs • Trade routes

Therefore, the unit you will use as a cluster should depend on what groups the elements you are interested in sampling are involved in. The box above illustrates possible clusters for various sample elements.

Step 2: Develop the sampling frame. In cluster sampling, a *sampling frame* is a list of clusters. In some cases, a list of the clusters you have defined may already be available, although you should make sure that the list is complete and up-to-date. Or, you may have to create a sampling frame. To do this, you might need to visit communities within your program's target area to make lists of schools, health facilities or youth organizations. You might also need to prepare sketch maps of villages, city blocks or neighborhoods to use as clusters.

Note

Avoiding bias

Bias is the extent to which the estimate of an indicator found by a survey differs from its "true" value. To avoid bias, sampling frames must cover the entire target population for a program's evaluation effort. For example, if you included only active program participants in your sample, you might get a biased view of the strengths of your program. By not sampling young people who dropped out of the program, you would miss some important information, for instance, the reasons some youth did not like the intervention.

Most programs have a list of affiliated clinics. If you are using the national statistics office's CEAs as clusters for conducting household surveys of youth, that list should be available. If not, you will have to conduct some preliminary fieldwork to develop a list of city blocks or neighborhoods.

If creating a list of all elements within the entire target population is not financially or logistically feasible, you can restrict your sample to a part of the target population. However, your evaluation findings will only be valid for the portion of the target population covered in your sampling frame. For example, it might be feasible for logistical reasons to include only clinics located in the central part of the city. A sample selected from this restricted sample frame would result in evaluation findings that would only pertain to central-city clinics. Conclusions about program effects for the city as a whole could be made only under the strong assumption that other clinics were as effective as those in the central city.

Step 3: Select a sample of clusters. Once you have developed your sampling frame, the next step is to select a sample of clusters by using either *simple random* or *systematic* sampling. For the household survey, for example, you might decide to choose a sample of CEAs using systematic sampling. The specific steps involved in systematic sampling are described in Appendix 1.

Determining which method is best to use will depend on the numbers of clusters in the sampling frame, and on some statistical considerations discussed later in this chapter and in Appendix 1. The question of how many clusters should be chosen is also discussed in Appendix 1.

Step 4: Select a sample of elements from the clusters you have chosen. Finally, you will need to choose a sample of elements from the clusters you have selected. This is often done using either simple random or systematic sampling, but you could use other sampling schemes, depending on the context. For example, after a sample of clinics is selected, you might conduct exit interviews of all young adult clients who appear at sample clinics on a randomly chosen day. Guidelines for which sampling methods to use in various scenarios are provided in Appendix 1. The following chart summarizes how the basic steps in cluster sampling might be applied to household, school and health facility surveys.

Cluster Sampling Schemes for Household, School and Health Facility Surveys							
Type of Survey	**Target Population**	**Cluster Definition**	**Sampling Frame (clusters)**	**Cluster Sample Selection**	**Sampling Frame (sample elements)**	**Sample Element Selection**	**Further Explanation**
Household	Youth, parents	Geographic area with fixed boundary: CEAs, city blocks, villages, etc.	List of CEAs or other geographic areas	Use systematic sampling to select CEAs (or other geo-graphic areas) from which you will sample	List of households within a chosen cluster	Use simple random or systematic sampling to choose an equal number of households from each sample cluster	See Appendix 1 for alternative sampling methods See Appendix 2
	Community leaders	Same as above	Same as above	Same as above	List of community leaders within a chosen cluster	Use simple random sampling to choose an equal number of leaders from each sample cluster, or select a sample of clusters and include all community leaders from those clusters	
School	Students, parents	Schools	List of all schools reached by the program	Use systematic sampling to select schools from which you will sample	List of all students within each school, or list of all classes within each school	Use simple random or systematic sampling to choose sample students from a list of all students, or choose sample classes and include all students	See Appendix 2

Cluster Sampling Schemes for Household, School and Health Facility Surveys							
Type of Survey	**Target Population**	**Cluster Definition**	**Sampling Frame (clusters)**	**Cluster Sample Selection**	**Sampling Frame (sample elements)**	**Sample Element Selection**	**Further Explanation**
School (continued)	Teachers, adminis-trators	Same as above	Same as above	Same as above	List of all teachers within each school, or list of classes within each school	Use simple random or systematic sampling to choose teachers from a list of all teachers, or choose teachers associated with the sample classes chosen for student surveys (linked strategy)	
Health facility	Service providers	Health facilities	List of all facilities reached by the program	Use systematic sampling to select facilities from which you will sample	List of service providers at each sample facility	Use simple random or systematic sampling to choose providers from a list of all providers at each facility, or interview all providers present on a randomly chosen day that a sample facility is visited	

Key Issues in Cluster Sampling

A number of issues will affect the way in which you choose samples:

- deciding whether to sample clusters or elements, or include them all;
- determining how many clusters to choose in a sample;
- sampling clusters of varying sizes;
- selecting different sample elements for the same survey;
- sampling for sub-groups; and
- deciding when to use general population household surveys to evaluate school-based or facility-based programs.

Decide whether to sample clusters or elements, or include them all.
There are some situations in which you may not need to sample clusters or elements. Deciding whether to sample clusters or sample elements depends on two things:

- **How many clusters you have:** For small-scale programs covering only a handful of program sites (for example, communities, schools or facilities), you may be able to include all sites in the evaluation.
- **How many elements exist within each cluster:** If the number of elements at each site is not too high, it may be simpler logistically to include all sample cluster elements in the sample. For example, if self-administered questionnaires are to be used in evaluating a school-based program (and if resources permit), it might be easier to include all classes in sample schools or all students in sample classes. This is also true for:
  - — administrators and teachers at sample schools,
  - — community leaders, and
  - — managers and service providers at health facilities.

Determine how many clusters to choose in a sample.
When choosing clusters for a sample, select as many clusters (e.g., communities, schools or facilities) as your resources will permit. Generally, you should choose at least 30 clusters, especially for large programs. A sample of more clusters of smaller size is always preferable to one with fewer clusters of larger size. Choosing a smaller number of clusters will reduce the precision or reliability of your data, thus making it more difficult to detect real changes in indicators. For example, it is better to choose 30 schools with a population of 200 students each, than to choose 20 schools with a population of 300 students each. If you choose fewer than 20 clusters, there is little advantage in randomly selecting clusters, because the desirable statistical properties of probability sampling do not apply to samples of fewer than 20.

In such cases, a "judgmentally representative" sample of clusters, in which the distribution or profile of clusters on important characteristics (e.g., size, location, socioeconomic level of the student body or catchment area served) is similar to the distribution in the target population as a whole will usually be acceptable. To do this when choosing a sample of schools, you must make sure that large and small schools, schools serving well-off and poor students and schools in different parts of the geographic area covered by your program are included in the sample.

You can find further guidance on deciding on how many clusters to choose in Appendix 1.

There are some issues to consider when sampling clusters of varying sizes.
When sampling clusters of varying sizes (e.g., schools with different numbers of students, geographic areas with varying population sizes), use systematic sampling with probability-proportional-to-size (PPS). PPS allows larger clusters to be given a greater chance of selection than smaller clusters. This compensates for elements in large clusters having a lower chance of being chosen at the second stage of sample selection than elements in smaller clusters.

Cluster sampling is the most widely used type of probability sampling in conducting surveys for program evaluation.

In order to do systematic sampling with PPS, you need to have a *measure of size* (MOS) for the clusters you are using. An MOS is a count or estimate of the number of sample elements associated with each cluster. Exact counts are not necessary; MOS approximations are sufficient. If an MOS is not available, sample clusters may be chosen using systematic sampling with equal probability.

Examples of an MOS for each type of survey include the following:

- **Household:** number of households; estimated total population; estimated number of young adults in the geographic area
- **School:** number of students
- **Health facility**: number of young adult clients seen at health facilities

You can find detailed steps for selecting a sample of clusters using systematic sampling with PPS and with equal probability in Appendix 1.

There are two ways to select different sample elements (e.g., students, parents, teachers, community leaders) for the same survey.
For example, suppose that you are conducting a school survey that includes students, parents and teachers as the sample elements. You have two options:

- **Choose sample elements independently of one another:** This would require a separate—or *independent*—sampling process for students, teachers and parents.
- **Include all sample elements in the same sample:** In other words, choosing parents and teachers associated with the same sample of schools or classes chosen for student surveys. This is a *linked* sampling strategy.

The primary advantage of the linked sampling strategy is that it will limit data collection to the same set of clusters, resulting in lower data collection costs. This strategy also enables you to associate or relate indicators for different elements to one another during analysis. For example, if the teachers and parents chosen for a program evaluation were also the teachers and parents of students chosen for the sample, it is possible to determine whether

changes in outcome indicators for students were associated with attitudes and behaviors of their teachers and parents.

Note: The same options are valid for selecting sample elements such as households, young adults, parents or community leaders (see Appendix 1 for examples).

You may want to sample for sub-groups.
If you want to measure indicators for specific sub-groups within your target population (e.g., by sex, age, ethnic group, economic status, urban/rural/peri-urban), you will have to choose a separate sample for each sub-group, in order to ensure that an adequate sample size has been obtained for each sub-group. For example, you may want to know whether your program has been equally effective for male and female young adults. If so, the only way to ensure that you will have enough male and female subjects is to select separate samples. The process of dividing the target population into sub-groups and then taking a sample from each is known as *stratification.*

It's important to know when to use general population household surveys to evaluate programs.
Use these surveys to evaluate school-based or facility-based programs only if the proportion of the population reached by the program is at least 50 percent and preferably 75 percent. General population household surveys are not an efficient means of evaluation for programs that cover less than 40 percent of the general population of young adults, because such large numbers of households would have to be contacted in order to find a sufficient number of young adults who had been exposed to a program.

Determining Sample Size

Many ARH outcome indicators are measured at the level of individual persons or events, so that sample elements may include youth, parents, community leaders, teachers, health service providers, peer educators and service transactions. *Sample size* refers to the number of sample elements from which you will need to collect data in order for your evaluation findings to be statistically significant. The size of your sample depends on a number of factors, including:

- the indicators chosen,
- the baseline value of the indicators in the study population, and
- the amount of change you want to be able to measure accurately.

When measuring changes in indicators, you will determine how many individuals or events need to be surveyed in each round of data collection so that you can accurately measure changes in the indicators you have specified. This section addresses calculating sample sizes for measuring:

- changes in indicators involving one behavior,
- changes in indicators involving more than one behavior, and
- changes in indicators for aggregate units, such as schools, facilities and communities.

Three points must be made about sample size:

- **Measuring changes in selected indicators specified at the beginning of the program over time requires a larger sample size:** The sample size required to do this is larger than the sample size needed to measure an indicator at a given point in time.

Sample Size Table
Minimum Sample Needed to Measure Changes in Indicators Involving One Behavior

Starting Level of Indicator (P_1)	Level of Indicator You Hope to Achieve (P_2)	Required Sample Size (n)
.10	.20	438
.10	.25	216
.20	.30	638
.20	.35	300
.30	.40	775
.30	.45	353
.40	.50	843
.40	.55	376
.50	.60	843
.50	.65	368
.60	.70	775
.60	.75	330
.70	.80	638
.70	.85	267
.80	.90	438
.80	.95	163

Sample sizes shown assume significance and power of 90 percent, a design effect of 2.0, and two-tailed statistical tests. See Appendix 2 for discussion of these parameters.

Note

For scenarios not covered in the sample size table
A more complete table showing sample size requirements for different combinations of significance and power is in Appendix 2. This Appendix also includes a formula for calculating required sample size, and guidance on choosing parameters for using this formula.

- **This discussion assumes that your indicators will be measured as proportions:** In fact, most youth program evaluations use indicators measured as proportions.
- **This discussion does not distinguish between use of probability and non-probability sampling methods:** It is noted, however, that having an adequate sample size will not minimize the risk of bias in non-probability samples.

Calculating sample size for measuring changes in indicators involving one behavior depends on the following five factors:

- The *starting level* of the indicator (that is, the level of an indicator at the time of a baseline survey). We will label this P_1.
- The *magnitude of change* in an indicator that you want to reliably detect. We will label this $(P_2 - P_1)$, where P_2 is the level of an indicator at the second round of data collection.
- The probability with which you want to be certain that the magnitude of change $(P_2 - P_1)$ did not occur by chance. This is referred to as the *level of significance.*
- The probability with which you want to be certain of detecting the magnitude of change $(P_2 - P_1)$, if one actually occurred. This is referred to as *power.*
- The proportion of sample elements in the target population that have the characteristics specified in the indicator being measured.

To help determine the sample size you will need to measure changes in indicators involving one behavior, see the table at left. The sample sizes listed are based on 90 percent significance and power.

To use this table:

- **Choose an initial or starting level of an indicator (P_1):** For example, your indicator may be the proportion of young adults who intend to use condoms in subsequent sexual encounters. Before you undertake your first survey, your best guess is that about 30 percent of young adults would intend to use condoms. Therefore, P_1 is set at .30.

Note: If it is not possible to estimate the starting level of an indicator, the safest course is to choose a value of P_1 equal to .50, because this will ensure an adequate sample size even if the true value of P_1 is different from .50.

- **Specify the magnitude of change you want to be able to detect reliably ($P_2 - P_1$):** For example, suppose that your program would be deemed successful if the proportion of youth that intend to use condoms was increased by 15 percentage points (to 45 percent). Therefore, P_2 is set at .45.

Note: Only changes of 10 and 15 percentage points are provided in the table. Measuring smaller changes requires larger sample sizes because greater precision is needed to measure small changes. Changes of less than 10 percentage points require a sample size that may exceed the resources available to your program evaluation effort.

- **Read the required sample size on the table:** With P_1 set at .30 and P_2 set at .45, a sample size of 353 young adults per data collection round would be required, according to the table.

Note: Remember that each round of data collection will require this sample size. In the above example, you would need a sample of 353 young adults in the baseline survey and another 353 in a follow-up survey.

Example

Calculating sample size

Suppose that a program's objective is to reduce the proportion of male students who believe that it is acceptable for male students to coerce female students into having sex. It is estimated that 50 percent of male students believe that this behavior is acceptable at the outset of the program (P_1 = .50). The program aspires to reduce this percentage by at least 10 percentage points (P_2 = .40). If 90-percent levels of significance and power are to be used, the required sample size is 843 male students per data collection round. This example illustrates the point that the sample size required to measure a change from .50 to .40 is the same as that required to measure a change between .40 and .50. This applies to equivalent changes between other levels as well.

Many evaluators opt to add a cushion for non-response and dropouts.

In every study, some subjects do not cooperate or drop out. This problem is usually addressed by increasing the target sample size by a fixed proportion (about 10 percent). If the chart above recommends a sample of 843, you may want to increase it by 10 percent to 927 to compensate for any non-response.

Dropout rates of 25 to 30 percent or more are common in follow-up studies of youth. To continue the example above, after having added a 10 percent cushion for non-response, you might consider adding another 25 percent to the sample size as a precaution against dropouts, raising the baseline sample from 927 to 1,159. This will help ensure that you will be able to detect changes in indicators if they actually occur, even if you encounter non-response and dropouts.

Calculating Sample Size to Measure Changes in Two Behaviors

Suppose a program objective is to increase the proportion of secondary school students who use a condom during their first sexual encounter.

- ➤ **Calculate the sample size you will need to measure the magnitude of change in condom use:** You may estimate that only about 10 percent of students currently use a condom during their first sexual encounter (P_1 = .10). Your program aims to at least double this to 20 percent (P_2 = .20). With 90 percent significance and power, the required sample size according to the table is n = 438 students per data collection round.
- ➤ **Calculate how many youth you will need to survey in order to get the sample size you need of youth who have had a first sexual encounter:** If current estimates are that 10 to 15 percent of students engage in their first sexual encounter each year, over a three-year project period, 30 to 45 percent of students would have their first sexual encounter. Since as few as 30 percent of students may actually go on to have their first sexual encounter during the three-year project period, the required sample size per data collection round is n = 438/.30 = 1,460 students.

Calculating sample size for measuring changes in indicators involving more than one behavior involves a second step.

For example, your program's indicator might be "Proportion of young adults who used a condom during their last sexual encounter." This indicator requires measurement of two behaviors: (1) the proportion of youth who have had a sexual encounter, and (2) the proportion of youth who used a condom during that encounter.

- ➤ **Calculate the sample size you will need to measure the magnitude of change in condom use:** Using the sample size table, determine how many youth would need to be surveyed in order to measure a change in the proportion that used a condom during their last sexual encounter. If you estimated that condom use at last sexual encounter was 30 percent (P_1) and you hope to increase this proportion to 45 percent (P_2), then you would need to sample 353 youth.
- ➤ **Calculate how many youth you will need to survey in order to get the sample size you need of sexually active youth:** This indicator can only be measured with youth that are sexually active, for example, those youth that had a sexual encounter during the last six months. Assuming that not all youth are sexually active, you will need to estimate how many youth would have to be interviewed in order to find 353 youth that are sexually active.

First, estimate the expected proportion of youth that had a sexual encounter in the last six months; for example, you may think that 40 percent of youth in the target audience are sexually active. Then, divide the required sample size (353) by the expected proportion of eligible respondents (.40). The number of youth you would need to interview in order to find 353 youth who are sexually active is (353/.40), so n = 883. Or, in other words, you must interview 883 youth in order to measure proportion who used a condom at their last sexual encounter.

The difficult part of calculating sample size for indicators involving more than one behavior is estimating what the appropriate underlying proportion is, for example, estimating what proportion of young adults are sexually active. You might base your estimates on other surveys or on anecdotal information from the field. If resources permit, however, conducting a small pilot survey can help you estimate this proportion.

If you are unsure what to estimate for a particular behavior, it is better to err toward *underestimating* the proportion engaging in the given behavior. The more you underestimate the given behavior in your target population, the higher your resulting

sample size will be. For example, if you are uncertain whether 40 or 50 percent of students in your target population typically engage in sex during any given six-month period, you should use the 40 percent figure to determine your sample size requirements.

Measuring changes in indicators for aggregate units requires a smaller sample.
Many projects may be interested in measuring changes in indicators at the school, facility or community level. Sample size requirements for measuring these indicators will be *one-half the size* of those shown in the sample size table. This has to do with design effects in cluster sampling, which are discussed in Appendix 1.

For example, suppose you wished to measure changes in the proportion of schools that had implemented life skills education programs that met certain guidelines. If you thought that only 10 percent of schools had such programs at the beginning of the observation period and you wanted to be able to measure a change of 15 percentage points, the sample size table shows that you would need 216 schools in the sample. Because you are sampling an aggregate unit, a sample only half as large would be needed—that is, only 108 schools (108 = 216/2).

The required sample size may exceed the number of available "settings."
Some programs may calculate the sample size required to evaluate a school-based program and find that the recommended sample size is larger than the total number of schools covered by the program. In this event, as many settings as possible should be covered by the evaluation effort.

Sample Size Computation

	P_1	P_2	n
Boys	.20	.35	300
Girls	.30	.45	353

Commonly Asked Questions About Sampling

My program has many indicators that will be measured through the same data collection effort. Which indicators should I use to estimate my sample size?
Ideally, you would calculate the sample size requirement for each of your indicators and choose the largest of these as the sample size to be used. This will ensure that the requirements of all indicators are satisfied. However, if you have a large number of indicators, this can be tedious. In this case, you can:

- ➤ select a few of your more important indicators,
- ➤ calculate the sample size requirements for each of them, and
- ➤ use the largest of these calculations as your sample size.

Keep in mind that this strategy has two drawbacks: (1) Changes in some indicators may be measured more precisely than needed, and (2) changes in other indicators may not be measured precisely enough.

What if I want to compare different sub-groups of young adults, schools or facilities?
Comparing changes in indicators for different sub-groups requires larger sample sizes. For example, if you want to measure changes in indicators separately for male and female young adults, you need to calculate the sample size separately for each gender.

For example, let's return to the program objective "Increase the proportion of secondary school students who intend to use a condom during their first sexual encounter." If you estimated that 20 percent of the boys intended to use a condom during their next sexual encounter (P_1), and you aimed to raise this by 15 percent to 35 percent (P_2), you would need a sample of 300 boys. If you thought 30 percent of girls intended to use a condom during their next sexual encounter (P_1), and you aimed to raise this to 45 percent (P_2), you would need a sample of 353 girls. Therefore, you would need to sample 653 youth altogether, 300 boys and 353 girls.

What if I plan to conduct more than two rounds of data collection?

The sample sizes discussed in the previous section pertain to the magnitude of change on indicators to be detected between any two rounds of data collection. If you are planning to undertake more than two rounds of data collection, you may be able to use smaller samples for two reasons:

- If conducting several rounds of data collection results in a longer period of observation between the initial and final round of data collection, larger changes in indicators are more likely, thus requiring smaller samples, and
- you may be able to use certain statistical procedures that require smaller samples (for example, repeated measures methods) in your analysis.

If you are planning more than two rounds of data collection, it is recommended that you consult with experienced researchers (for example, at a local university) on these issues before deciding on what sample size to use.

Should I retain the same sites in each round of data collection?

From a statistical point of view, the advantage of going back to the same sample of schools, facilities or communities in each round of data collection is that it will generally increase the precision of your evaluation. This is especially true when relatively small numbers of sites are being covered. One potential drawback, however, is that going back to the same sites may cause the program to be implemented more rigorously in the sites chosen for the program evaluation than in other sites, therefore resulting in bias. While retaining the same sample of sites over the course of an evaluation is generally recommended, you as a program manager must ensure that the program is not implemented more rigorously at sites chosen for the evaluation than at other program sites not chosen for the evaluation. If that should happen, your evaluation results will be biased.

Should I retain the same sample of young people in each round of data collection?

Evaluations in which the same young people are covered in each round of data collection are called *panel studies*. For the same reasons that retaining the same sample of schools, communities, etc. increases the precision with which changes are measured, panel studies are the preferred way to measure changes in behaviors among young people. However, panel studies depend on retention of the subjects, which is time-consuming and costly. Accordingly, panel studies conducted over extended periods of time (i.e., two or more years) are recommended only when your program has a long-term evaluation and substantial budget.

Note, however, that collecting data from the same individuals over short periods of time is often feasible. For example, many evaluations of school-based programs follow samples of students over one- or two-year periods.

DATA COLLECTION AND THE M&E WORKPLAN

CHAPTER AT A GLANCE

- ➤ Reviews data collection steps
- ➤ Addresses ethical concerns
- ➤ Presents options for data collection methods
- ➤ Discusses tasks involved in developing an M&E workplan

Preparing for Data Collection[1]

How you prepare for data collection will influence the quality of the data you collect. If you or your staff collect the data, you will need to:

- ➤ address any ethical concerns,
- ➤ prepare written guidelines for how data collection will be done,
- ➤ pre-test data collection indicators, instruments and procedures, and
- ➤ train all staff who will collect the data.

Address ethical concerns before planning for data collection.
Ethics refers to both professional standards of conduct as well as moral principles and values exercised in conducting research and evaluation studies. Ethical reviews are designed to consider and mediate the potential risks and negative consequences to participants as a result of their participation in a study or evaluation. Most programs are already addressing ethical concerns in the implementation of their interventions. For evaluators, responding to ethical concerns will influence your relationship with the community and enhance your ability to collect quality data. The more ethical your data collection effort, the more honest and reliable the information you collect, ensuring that your M&E results are valid.

Ethical concerns can be addressed in many ways:[2]

- ➤ **Community input:** The M&E effort should respond to concerns of community stakeholders, such as parents groups, youth clubs, NGOs, religious groups and youth. Many may have strong views on what kinds of questions are relevant and acceptable. For example, some youth may voice their concerns about participating in a survey that explicitly addresses sexuality. Others may have suggestions about developing

[1] Many of the ideas presented here about how to overcome the challenges of collecting data from youth are based on discussion from the *YARH Measurement Meeting*, September 28–30, 1999. In particular, presentations by Gary Lewis, Johns Hopkins University Center for Communications Programs, and Paul Stupp, Centers for Disease Control Division of Reproductive Health, were helpful in developing the substance of this chapter.

[2] Based in part on a presentation by Cynthia Waszak, Family Health International, made at the *YARH Measurement Meeting,* September 28–30, 1999.

questionnaires that young people will feel comfortable responding to, thereby increasing the validity of their responses. Document community input for reference in case the evaluation becomes controversial.

Applying ethical standards in data collection improves the quality of data.

- **Parental permission:** Local standards and laws will determine the age at which a young person can consent to answer a survey or questionnaire. For example, in some communities, a young person is considered an adult at the age of 16; in others, "emancipated minors" include only those who are married or serving in the military. Be sure to find out when you need to secure parental permission. If no legal standards exist, ask for advice and input from the community. Consider whether it is practical to obtain permission from parents, and explore options for getting "adult advocates" to permit young people to participate in the evaluation. Parental consent can be verbal, but written is best.
- **Informed consent:** Before completing a questionnaire, youth must understand what they are being asked to do and how the information they provide will be used. This is called *informed consent.* Informed consent is key to getting good data, because youth may not answer questions honestly if they are concerned about who will see their answers or how the data will be used. Most questionnaires have an informed consent waiver that is read to youth before they begin the survey. The language of informed consent must be explicit and comprehensible to young people. Those collecting data should be trained to understand the concept of informed consent, and able to answer concerns from youth about their participation. Finally, you should consider legal requirements to disclose certain types of information, such as sexual abuse and illegal drug use. If you are required by law to disclose this information, you must explain this to participants before they answer any related questions.
- **Voluntary participation:** You must ensure that young people are answering your questionnaire voluntarily. Compensation is acceptable, if it is modest (e.g., reimbursement for transportation costs, snacks). Avoid pressure by authority figures to participate in a survey conducted in an institutional setting. If possible, identify adults who can advocate for youth, such as a school nurse, and answer the questions of youth who are concerned about participating in the evaluation. Senior-level supervision can help ensure that the principles of voluntary participation are followed in the field. You should also develop a system to minimize how peers influence their friends' decisions to participate in the evaluation.
- **Confidentiality and privacy:** Most evaluations of youth programs collect information anonymously. If names are collected, keep them separate from completed questionnaires and connect the two by a code. Privacy is also important. The survey should be

administered in a private place, where others cannot hear or see a young person's answers. Avoid testing techniques that inadvertently reveal something about youth (e.g., longer questionnaires for sexually active youth).

- **Risk to respondents:** Even with confidential surveys, concerns remain that certain kinds of questions are harmful to youth. Questions about sexual abuse, for example, can deeply upset youth who have been abused but have never talked about it before. Some evaluators opt not to ask these kinds of questions unless there is a service provider available. Many evaluators develop a skip pattern in their questionnaires so that youth who have never heard of sex, for example, are not asked about sexual behavior. Some evaluators offer a discussion or question-and-answer period after youth have completed their questionnaires.

Prepare written guidelines for data collection to help ensure high-quality data.
High-quality data collection efforts usually have detailed written instructions on how data will be collected. For large-scale surveys, such as country-level Demographic and Health Surveys (DHS), detailed manuals are prepared as guidelines for interviewers and field supervisors. Guidelines provide step-by-step instructions, as well as guidance on handling problems and questions that might arise during the course of data collection.

While your program may not have the resources to prepare manuals comparable to those used for the DHS, you should develop guidelines for each of the data collection instruments you plan to use. Guidelines ensure some degree of standardization in the data collection process. Without guidelines, each person collecting data will use his or her own method, procedure and problem-solving strategy. This may produce random error, sometimes referred to as *noise*, in the data. Without standard procedures, evaluation data may be biased because they were not all collected in exactly the same way.

Carefully design and pilot-test survey questions.
While designing a survey instrument, use qualitative exercises to explore young people's meanings and perceptions about the issues you want to measure. Young people's language, priorities and behavioral concepts are different from adults'. For example, a young person might think that taking a single contraceptive pill before having intercourse will protect against pregnancy, and thus will report having used birth control at last intercourse. Also, words can be interpreted in many ways; for example, "friend" can mean many different things to young people.

Budget enough time and resources to pilot-test the questions you plan to ask youth. Use the data collection instruments in conditions that are as similar as possible to your expected field conditions. For example, pilot-test materials with respondents who have the same socio-demographic characteristics of the respondents you plan to collect data from, but in areas where the program is not conducted (to protect future results). Pilot tests will:

- detect questions that respondents have trouble understanding,
- verify how long it will take to collect the data,
- build competence in data collectors, and
- uncover problems in field procedures.

Even if your organization is experienced in data collection, conduct a pilot test. It is best to discover and correct any problems before data collection begins.

Order topics on the survey so that sensitive questions are asked last.
In survey instrument design, questions should move from simpler issues to more complex ones. At the same time, surveys should ask less sensitive questions first, followed by more sensitive ones. What youth consider "sensitive" varies from place to place. For some, reporting about the number of pregnancies is more sensitive than answering questions about premarital sex. In other areas, the most sensitive questions will be about economic issues or partner violence. One way evaluators have increased young people's comfort in answering questions about sexual behavior is to ask about marital status only at the end of the survey. Asking sensitive questions at the end of a survey also ensures that you will have collected most of the information you're interested in, even if a young person refuses to answer the sensitive questions.

Build in checks to measure the validity of responses to sensitive questions.
All evaluators are concerned about accuracy of responses, especially on accounts of sexual experience. To make sure the data are valid, ask the same question—in several different ways—throughout the survey, and then compare the responses. For example, ask the date of first sex, the date of marriage and whether the person had premarital sex. If answers contradict one another, you may want to drop that case from the analysis of that particular issue. You might also want to conduct an anonymous, sealed, written questionnaire on a sensitive question and determine whether there are contradictions. Spot-check surveys, paying close attention to responses on sensitive questions, and discuss any problems with data collectors.

Asking about young people's perceptions of their peers' behavior may also provide data against which to check the validity of responses to sensitive questions, although in some cases youth perceive that their peers are taking more risks than their peers themselves report doing.

Set a reasonable length for the data collection instrument.
Be sure that you are able to collect the information you really need before youth become bored and refuse to participate in the survey. Collecting too much data will also cause problems in analysis; many evaluators collect so much information that they never analyze all of it fully.

Select data collectors that youth will respond to.
Who collects information from youth may be the most important factor in the validity of your data. In some communities, older, mature interviewers work better; in others, younger interviewers appear to be more effective. In almost all contexts, the age and sex of data collectors are the two major factors to consider. Conduct focus groups with youth before hiring data collectors to find out whom they feel most comfortable with. You might have to balance youth's preference for younger data collectors and your own need for a well-educated team.

Train every person who is collecting data.
Staff, youth program participants or professional interviewers may be involved in data collection. Regardless of what experience data collectors have, training should include:

- an introduction to the evaluation objectives,
- a review of data collection techniques,
- a thorough review of the data collection items and instruments,

- ➤ practice in the use of the instruments,
- ➤ skill-building exercises on interviewing and interpersonal communication, and
- ➤ a discussion of ethical issues.

Of these, training on interpersonal communication skills—such as establishing a comfortable rapport, maintaining privacy and confidentiality and treating the subject respectfully—is essential when collecting data from youth. Role playing, in which people practice collecting data, can be a useful training device. If possible, training should be concluded with a practice data collection exercise at a site that will not be part of the evaluation sample. After this practice, participants should review their experiences with the entire group.

It is important that data collectors are not a source of unreliability. It is therefore necessary to train them in correct use of the instruments. After the training, administer competency exams and discuss any remaining deficiencies.

A sample training schedule is provided at right. This example represents the minimum time to spend in a training exercise. The length of training you need depends on two things:

- ➤ the number and complexity of data collection instruments to be used, and
- ➤ the experience of the persons being trained with these instruments.

In many cases, it is more appropriate to hire outsiders to collect data. For example:

- ➤ If staff conduct exit interviews with clients, clients may give less than candid assessments of program services (this is referred to as courtesy bias).

Sample Training Schedule	
Day 1	• Introductions • Review evaluation purpose and objectives • Discuss possible problems with data collection • Discuss sample selection • Provide overview of data collection instruments • Review data collection instruments
Day 2	• Provide guidelines for conducting interviews • Divide into groups; role-play interviews • Discuss experience of role playing
Day 3	• Practice data collection at selected sites • Review experiences practicing data collection • Discuss how to resolve problems encountered during data collection • Role-play corrected procedures
Day 4	• Plan travel logistics • Cover administrative and financial details

Adapted from Miller et al., 1997.

Note

Guidelines for managing an outside group to undertake data collection

- Develop a written contract outlining exactly what is to be done, by when and how much it will cost.
- Specify the use of data collection methods and instruments, and provide instructions in their use.
- Insert provisions to ensure adequate data quality (such as a minimum number of field supervisors) and procedures to minimize data entry errors.
- Include pilot-testing of data collection instrument, with findings communicated to a designated staff person.
- Maintain final approval of all procedures and data collection instruments before they are used.
- Insist that changes made to procedures or data collection instruments be approved by a designated staff person.
- Designate a knowledgeable person from your staff to serve as a liaison with the contractor to answer questions and monitor data collection work.
- Ensure adherence to the sampling plan.

- Program staff who are needed full-time to implement the project will often not be able to spare the time to collect data.

In such cases, the preferred course of action is to hire an outside group to collect data for you. While hiring outsiders will reduce the amount of preparation needed, you will still have to manage the outside consultants.

Focus your data collection efforts on key issues and ensure that all interviewers are adequately trained.

Limit the length of the entire data collection exercise.
You may find yourself facing a trade-off between the number of interviewers you have to collect information and the time you allow for fieldwork. In some cases, it might be more manageable to collect information over a longer period of time with a smaller group of data collectors. However, some people have found that data collectors get tired of interviewing youth after 2–3 months and may drop out if fieldwork lasts too long.

Types of Data Collection Methods

Monitoring and process evaluation data are collected using a variety of methods, such as reviewing service statistics and administrative reports, and conducting surveys, focus groups and interviews. Once you have selected a data collection method, you will need to develop appropriate *instruments*. An instrument is the tool you will use to collect information, such as a form that your staff completes or a survey in which youth participate.

Monitoring data allow you to track whether your program is reaching its desired audience.
Most monitoring consists of tracking and counting activities related to program implementation. Some data for program monitoring indicate whether systems are being implemented as planned; for example, your program may record how many volunteer trainings took place over a one-year period. Other data for program monitoring are recorded in conjunction with service delivery. Basic service statistics about program operations (such as client visits or trainings offered) can be found in event logs, registers or tally sheets.[3] These data are often communicated to program managers via monthly or quarterly reporting forms.[4] Supervisory or administrative reports, which are often narratives, provide program managers with more in-depth insights into problems being encountered in program implementation and possible solutions.

Don't go overboard collecting too much data. Focus your efforts on key areas so that monitoring does not detract from program implementation.

[3] See Instrument 2 for examples of tally sheets.

[4] See Instrument 3 for examples of reporting forms.

Process evaluation data allow you to observe and interpret how your program is working.
Process evaluations help you assess the quality of your program's staff and volunteer performance, the quality of the activities and events you undertake and how staff, participants and the community are reacting to the program.

Developing checklists

Checklists, which enumerate key features of a setting or process, serve two useful purposes:

- They guide observers in covering key process evaluation topics in a systematic manner.
- They establish the criteria or standards to be used in assessing how program activities are being carried out.

For example, in assessing a staff training program, a checklist might be developed that enumerates the key features of the training setting (such as adequacy of space, lighting, tables and chairs and necessary audiovisual equipment) and the standards for trainer performance (such as organization, responsiveness to questions and communication skills). Several examples of checklists are provided in Instrument 1.

Inventorying facilities and services

A facility inventory form can assess the adequacy of facilities and equipment and verify that:

- the facility meets program standards in terms of space, lighting, arrangement to ensure privacy, etc.;
- the equipment needed to provide program services is on hand and in working order; and
- the facility has adequate supplies of expendable items to satisfy demand for services.

Sample inventory forms are provided in Instrument 6.

Interviewing key informants

An individual in-depth interview with "informants" can produce valuable information. Individual in-depth interviews are preferable to focus group discussions (where people are influenced by group dynamics) and are often more practical.

The process of interviewing can range from highly structured to completely open-ended. Highly structured interviews use a questionnaire (such as the Interview Guide for Staff Providing RH Services in Instrument 8) that has pre-coded questions and answers that evaluators use to interview respondents. You would use a highly structured interview when you want data that can be standardized, or if you want to collect data that are easily managed.

Semi-structured interviews use a questionnaire that has a mix of structured and open-ended questions, such as the Questionnaire for Debriefing Mystery Clients in Instrument 10. A semi-structured interview allows you to ask open-ended questions that will help you to determine the perspectives of your program participants and staff.

With unstructured or open-ended interviews, an evaluator may have a list of guiding questions to start the discussion but will allow the respondents themselves to shape the discussion. Unstructured interviews are useful because they allow respondents to identify and discuss the issues that they think are most important, rather than being guided by the interviewer's questions. However, they can be difficult to combine, code and analyze because of different content, depth and duration.

In-depth interviews can help you understand young people's actions and

behaviors and how the youth interpret their own actions. For example, you might ask youth who have had STIs how they think they got the infection, whether they sought advice from peers and/or important adults in their lives and where they received treatment. You might also ask them what would happen if their parents or extended family members found out they had an STI, or what they believe the general community thinks about youth who get STIs. In-depth interviews may also allow you to explore new issues or uncover issues you had not realized existed or were relevant. For example, a young person may tell you about a home method of treating an STI that you had never heard of before, or express a concern about a side effect of STIs that you had not realized existed.

Surveying staff

Staff surveys are a cost-effective way to gather comparable information from a large group of people. Surveys may be self-administered (completed by the respondent with a pencil and paper), or may be administered by an interviewer.

These surveys can assess technical competence, attitudes toward providing services to adolescents, perceptions of service needs, and knowledge or mastery of topics and skills learned in training. This type of survey is often conducted in conjunction with observations of service transactions and interviews of program participants or clients in order to assess the quality of program services. A sample of a staff interview guide is provided in Instrument 8.

Conducting exit interviews with clients

Exit interviews are conducted after clients have participated in a program activity or received a program service. Interviewers can ask about:

- how clients thought they were treated by service providers,
- how long they had to wait,
- whether they received the service or services they came for,
- their assessment of the services and the facility, and
- whether the service provider gave them enough information about the service provided and any follow-up steps they need to take.

This approach can be used with clients of health services, peer education efforts, youth center programs, workplace programs, etc. Like interviews with key informants, client exit interviews can be either highly structured, based on a pre-coded questionnaire, or unstructured, with open-ended questions.

An example of a client exit interview questionnaire is provided in Instrument 9.

Using mystery clients

The *mystery* or *simulated client* data collection strategy entails sending trained persons (including trained adolescents) to program facilities in the assumed role of clients, who then report on their experience. For example, an adolescent mystery client might be sent to health facilities to seek counseling or contraceptive services. Afterward, he or she

Note

Ethical concerns around using mystery clients

Some observers believe that undisclosed observations or mystery client visits are unethical because they involve misrepresentation. However, others believe that the benefits outweigh those concerns. They say that because the purpose of using mystery clients is to assess and improve the quality of services, it serves the interests of both clients and the program. In some ways, they say, "mystery" visits can be viewed as substitutes for supervisory visits.

would either complete a questionnaire or be interviewed. The two primary reasons to use the mystery client approach are:

- to avoid the bias in the service delivery process that often results from having service transactions observed, and
- to gather a sufficient number of observations of service transactions when the actual volume of service visits by adolescents is low.

Several examples of mystery client scenarios and an illustrative debriefing questionnaire are provided in Instrument 10.

Conducting observations

While interviewing helps you learn about people's attitudes and values and what they think or say they do, direct observation allows you to witness what people actually do and how they act in particular situations. For example, you can observe how a pharmacist reacts to youth who are seeking treatment for STIs, or what messages a peer educator gives to a young person during a counseling session.

There are two types of direct observation: *obvious* and *unobtrusive.* If your observation is obvious, people know that you are watching them and may therefore demonstrate what they think you want to see. For example, if you observe a peer educator during a counseling session that is usually one-on-one, your presence is very likely to influence the dynamic of the situation. Because they are being observed, the peer educators may give messages that they would normally not give, or youth may not ask questions they would normally ask. If you are unobtrusive, the people being observed do not know that you are watching them. For example, you might pretend to be a customer in a pharmacy and listen to how the pharmacist reacts to a young person seeking treatment for an STI.

While this may eliminate the problem of the person being observed "reacting" to you, it presents ethical problems, such as lack of confidentiality and not having informed consent from those being observed.[5]

Soliciting unstructured feedback from clients

Program managers and staff should always seek informal opportunities to obtain feedback from clients about the program. Such information supplements more formal process evaluation activities and may call attention to unforeseen problems or issues.

Free-listing, pile sorting and ranking[6]

When working with youth, it is necessary to elicit and understand local terms and slang. *Free-listing* is a technique in which you elicit words used to refer to a particular subject of interest.[7] An interviewer begins with a particular topic and asks the respondents to list the terms that correspond to that topic. For example, an interviewer may ask youth to list the symptoms and names of STIs. The interviewer can then ask questions to get more information about the meaning of each term. This will position the interviewer to use the language that youth are familiar with when asking questions about STIs on a structured survey questionnaire.

Pile sorting, ranking and scoring are techniques used to organize the terms generated from free-listing. In *pile sorting,* informants are asked to write the terms generated by free-listing on cards. Informants then sort the cards into piles so that each pile consists of items that are considered similar to one another.[8] The piles can be defined in any way, and

[5] Bernard, 1994.

[6] For a step-by-step guide on how to conduct free-listing, pile sorting and ranking, see Shah, Zamberi and Sumasky, 1999.

[7] Weller and Romney, 1990.

[8] Weller and Romney, 1990.

informants can be asked to sort cards into any number of defined piles. For example, if you do pile sorting after eliciting the names and symptoms of STIs, youth might sort cards into piles that reflect:

- the severity of each infection (each pile representing a different level of severity),
- whether they think males or females are more likely to become infected (separate piles for males and females), or
- how common they think each infection is among their peers (each pile representing those that are more common or less common).

Pile sorting allows you to determine what criteria are most important to youth, since the youth define the categories that each pile sort should be separated into. They can also help determine the similarity between certain terms when pile sorts are done with more than one group of youth. It is a very easy technique that requires relatively little time to administer and analyze.

Ranking and scoring techniques are also used to organize information generated by free-listing to analyze preferences, prevalence and decision-making processes. With *ranking,* respondents evaluate possible options and then present them in a sequence of preference or priority. For example, after free-listing and pile sorting names and symptoms of STIs, an interviewer might then have respondents free-list all the ways youth seek treatment for STIs. These options could then be ranked in terms of where youth most prefer to receive treatment.

With *scoring,* participants assign a value to each option rather than ranking it. Scoring allows for more in-depth analysis because it reveals the different "weight" assigned to each option. For example, youth could score the ways that they sought treatment of STIs. More frequently used options would receive higher scores, revealing preferences among youth for seeking services.

CONVENING FOCUS GROUPS OR INFORMAL LISTENING SESSIONS

Focus group discussions are used to identify issues, terms and interpretations from a group of individuals with similar characteristics. These discussions are often planned in advance, usually with 6 to 10 participants invited. The facilitator guides the discussion with lead questions and probes in order to gain an in-depth understanding of the attitudes, beliefs and perceptions of a specific group of people. Careful selection of questions will ensure useful information.

Informal listening sessions are a less structured and more spontaneous method of group discussion, for example, talking with peer educators after their monthly meeting. The focus should be on listening, with respondents bringing up and discussing topics of greatest importance to them.

Focus groups and informal listening sessions often provide insights about cultural norms. However, these group discussions usually do not reveal how individuals' opinions or behaviors deviate from those norms. It is therefore useful to supplement focus group discussions with in-depth interviews. Analyzing focus group discussions can present some challenges,which will be discussed further in Chapter 8. Instrument 13 provides a focus group discussion guide for use with in-school youth.

Mapping[9]

Mapping is the creation of a visual representation of key features of a residential area, work area or other area of relevance, such as your program's area of geographic coverage. Here we describe three different kinds of mapping: social, census and body mapping.

Social mapping indicates the boundaries of an area, its social infrastructure, its housing patterns or any of its other structures that are relevant to social interaction. Social maps may help you understand where particular groups gather and carry out certain activities, and may reveal diversity within a particular area. For example, a social map may show where youth gather and spend their free time, and allow you to assess whether peer educators are conducting outreach in the appropriate places.

Census mapping is used to gather information about a geographic area, including household data such as the number and gender of youth, education and literacy levels, employment and resource ownership. It is useful because it generates numerical data about the community and can provide more specific data, such as the number of youth out of school or involved in income-generating activities.

Body mapping entails youth drawing maps of the female and male bodies, focusing on the details of the reproductive system and how it functions. This method reveals participants' level of knowledge, the type of information young people have about the human reproductive system, and local explanations of reproductive and other health functions and terminology. For example, once youth draw a female body map, you might ask them how STIs affect the female reproductive system, where symptoms will appear and what parts of the reproductive system will be affected.

Qualitative methods like free-listing, group discussions, mapping, and case studies may provide valuable insights into the issues in young peoples' lives.

Preparing case studies, stories and portraits

Case studies, stories and portraits, written by evaluators, include stories or anecdotes heard during group discussions or interviews. They often describe a significant event in a person's life. They may also cover how a particular activity or service affected a participant's life or highlight particular problems, issues or program accomplishments. In general, case studies, stories and portraits add more "life" and meaning to evaluation findings and may provide valuable insight into the issues in people's lives.

[9] For a step-by-step guide on how to conduct mapping exercises, see Shah et al., 1999

Outcome and impact data allow you to measure the extent to which outcomes are achieved.
Conducting outcome and impact evaluations will require that you collect data using some of the techniques described above, as well as others.

POPULATION SURVEYS

Population surveys are highly structured surveys of a sample of the program's target population, as well as of youth who do not receive the intervention but live in areas that resemble your program's location.[10] Questions on the survey relate to the indicators of your program's outcomes, allowing you to determine whether your target population was exposed to program activities and experienced changes in outcomes. Instrument 12 contains a comprehensive questionnaire for evaluations of youth programs.

Note

Primary data are collected from youth specifically for the evaluation. *Secondary data* consist of information about youth available through recent surveys, already-existing qualitative research or administrative records.

COMMUNITY SURVEYS

Community surveys measure indicators related to outcomes at the community level by surveying stakeholders, youth and parents. The outcomes they address may include youth access to reproductive health information, adult communication with youth about reproductive health, and community support for youth programs. Questionnaires used for community surveys might be structured or open-ended. Instrument 11 is an example of a community questionnaire.

Selecting Appropriate Data Collection Methods

Your selection of methods will depend on your program needs and the population you serve. A few guidelines are suggested below.

Keep data collection simple.
Collect only essential data that can be analyzed and interpreted quickly, so that M&E results are provided in a reasonably short period of time to your program's staff, stakeholders and funding agencies.

Select methods based on availability of existing data and need for new data.
Consider the following issues as you select methods:

- availability of existing data,
- need for new data,
- capacity and availability of staff to help with data collection,
- need for outside assistance,
- timing of the data collection, and
- use of multiple data collection methods.

You should consider the availability of data for your M&E efforts prior to the start of your program's activities and services. This will give you time to assess available data and/or collect new information at the ideal time—your program's starting point. It may be possible to rely on recent surveys, qualitative community research or administrative records as your baseline data. However, even if these data exist, they may not be in a format that is useful for your purposes and will probably have to be extracted from current records or reports.

If you need new information, you should review the following points:

[10] Sampling is discussed in Chapter 6.

- ➤ Determine what is available from existing sources, e.g., client records or a survey.
- ➤ Identify what information is not available.
- ➤ Indicate what information is needed regularly and what is needed only periodically.
- ➤ Consider simple, inexpensive methods to collect information to supplement existing data.
- ➤ Review how information can be used by different groups (such as clinic staff, peer educators or media campaign workers) to help them in their work.[11]

Consider the capacity and time of staff and others for data collection.
Data collection for evaluations may involve your program staff, different stakeholders and sometimes outside researchers or evaluation experts, e.g., from a local university or funding agency. Who is actually involved in collecting data depends on who has the time and expertise to do it. You can best judge the capacity and time of your staff to collect data or to assist others with data collection.

You must first determine how frequently information should be collected for your various indicators. Assuming you are fortunate in having baseline data from the start of your program, the interval until follow-up data collection will need to be determined. The interval will depend both on the nature of the indicator and on the method of data collection. For example, assume you wish to assess whether the program has successfully upgraded youth centers to make them more attractive to young adults by adding new equipment. Approximately one month after the center is scheduled for its upgrade, you might complete a facility checklist during a site visit and interview the center director to determine whether equipment has been delivered and installed. In this case, data collection can take place immediately after the intervention. If, on the other hand, you wish to measure changes in youth's behavior that result from participation in peer education programs, data collection must take place six months to a year after exposure to the program. Data collection methods might include focus groups or surveys.

Advantages of Quantitative and Qualitative Methods[12]

Quantitative (e.g., surveys, records)	Qualitative (e.g., focus groups, case studies)
• Consistent and comparable data are collected. • This is a cost-effective way to collect data from large populations. • Self-administered survey questionnaires protect the privacy of respondents and may result in more honest replies. • Questionnaires administered by interviewers are suitable for obtaining data from people who are illiterate. • These methods ensure standardized data collection over time. • Data collection and sampling approaches enable findings to be generalized to larger populations. • Larger sample sizes can be used because data collection is usually less time-intensive.	• The views of young adults, parents and community members are obtained. • Social and cultural contexts experienced by young adults are defined. • Local slang used to describe sexual or reproductive health behavior is identified. • Questions for an interview questionnaire may be formulated and pre-tested. • Vocabularies for health education programs may be developed. • Messages for communication and medical campaigns may be pre-tested. • Unintended results of the project that might not be discovered through use of structured questionnaires may be revealed. • Program goals (e.g., empowerment, increased self-esteem, stronger negotiation skills) that are difficult to measure quantitatively may be assessed.

[11] Wolff et al., 1991

[12] Scrimshaw and Hurtado, 1987.

Combining Quantitative and Qualitative Methods	
Quantitative Method (Survey)	
Question: Did you use a condom the last time you had sexual intercourse?	Response: Twenty percent of adolescents who were sexually active said "yes."
Qualitative Method (Focus Group)	
Question: What are the reasons you don't use condoms?	Responses: • Using condoms creates distrust between partners. • Purchasing condoms is embarrassing. • You don't always know if you are going to have sexual intercourse or need a condom. • If you're a girl and you carry a condom, boys might think you are promiscuous. • I take (my girlfriend takes) birth control pills and we don't need to use condoms.

Use a variety of methods.
By collecting data in a variety of ways, you will get a more accurate picture of the progress and impact of your program. A review of service statistics, clinic records or surveys involves collecting *quantitative,* or numerical, data. These can be accompanied by *qualitative* methods such as focus group sessions, mapping or case studies, which involve descriptive or text information. Some data collection methods may be characterized as both quantitative and qualitative. For example, survey data typically yield quantitative information but may include open-ended questions that give more qualitative information.

The box above provides an example of how information from a qualitative data collection method (focus group) enhances the information from a quantitative method (survey).

When the information from data collection methods is combined, it shows not only what effect a program is having but also why it is having this effect. This more detailed information can then be used to modify the program's strategy.

Develop indicators and instruments that are sensitive to your program and population needs.
Whatever data collection method you use, the following process will ensure sensitivity:

- ➤ Identify local meanings, terms and issues using qualitative techniques, and identify specific issues for youth.
- ➤ Develop and decide on the most relevant issues, categories and terms.
- ➤ Quantify the most relevant issues and concerns.
- ➤ Analyze and interpret results.

Each of these steps is discussed below.

IDENTIFY LOCAL MEANINGS, TERMS AND ISSUES

In Chapter 2, we discuss the importance of defining and understanding the social and cultural contexts of the young people you are interested in reaching. In order to understand how a community characterizes adolescence, you might use group discussions to have community members categorize the stages of growth from childhood to adulthood. Their responses would be probed by facilitators until the local meaning of "adolescence" becomes clear. Similarly, youth may use specific terminology or slang to describe relationships, types of social groups, types of sexual partners, body parts, symptoms of illness or disease, sexual acts and other kinds of behaviors will help you understand how adolescence is experienced. It is important to identify these terms if they are categories you will measure as outcomes.

Data collection methods that help explore issues and identify local meanings, terms and issues include:

- ➤ focus group discussions,
- ➤ informal interviews,
- ➤ free-listing and pile sorting,
- ➤ semi-structured interviews,
- ➤ in-depth interviews, and
- ➤ observation.

In some cases, you may want to simply adapt terms and concepts for use in your outcome evaluation data collection instruments. Focus groups, free-listing and semi-structured interviewing are useful for such purposes. In other cases, particularly in process evaluation, you may want to explore a phenomenon in depth. For this you may consider a group discussion that includes participatory and visual analytic techniques, or in-depth interviews that allow you to probe deeper into individual perceptions, experiences and concerns.[13]

Develop and decide on the most relevant issues, categories and terms

Once you have conducted some exploratory data collection to identify terms and uncover issues, you will want to identify the issues that are of greatest relevance to your program and participants. Look for terms or concepts that appear frequently and emerge as patterns, as well as new issues that emerge that you want to explore further or quantify.

Analyze your data critically, considering what type of informant gives specific information and the conditions under which the information is provided. Be receptive to variations in responses and look for patterns that might explain these variations. You should also cross-check your data by using more than one method to collect information on the same topic, and by collecting information on the same topic from different groups of respondents.

[13] Shah et al. (1999) provide a step-by-step explanation of how visual analytic techniques are used with adolescents.

Quantify the most relevant issues and concerns

Once the most relevant issues have been identified, use this information to develop instruments that can collect consistent, comparable and quantifiable data. For example, you may have discovered that young people seek out a variety of reproductive health services, including traditional healers, pharmacies, vendors, private clinics and other sources. In a structured questionnaire, you could list each of these as responses to the question, "The last time you had a health problem, where did you seek treatment?" You will then be able to calculate the proportion of your total sample who sought treatment from each type of provider the last time they had a health problem.

Analyze and interpret results

Analyzing quantitative data consists of several steps that involve tabulation and statistical analysis. The steps involved in analyzing quantitative data are thoroughly discussed in Chapter 8.

Qualitative techniques, such as case studies and in-depth interviews, are often used after quantitative data analysis in order to help you interpret survey findings. For example, if there has been an increase in service use or a change in behaviors, you may want to ask selected key informants to explain the reasons they think those changes occurred. Similarly, qualitative methods can be used to assess program goals and outcomes that are difficult to quantify, such as stakeholder participation in the program. Qualitative techniques can help you determine participants' perceptions of the program, how they think it has changed their lives, what kinds of contributions participants made to the program, how they believe the program has impacted their community and what directions it should take in the future.

The table below and on the next 3 pages give a comprehensive list of possible data collection methods and sources. The final column of each Indicator Table at the end of Part I of this Guide refers to these illustrative data collection instruments. In Part II of this Guide, instructions are provided for preparing and carrying out several of the most common data collection methods.

Data Collection Methods: Potential Sources, Advantages and Disadvantages

Data Collection Method	Sources of Data	Advantages	Disadvantages
Review service statistics (clinic or outreach site)	• Client records • Family files • Clinic service register • Client logs All records can be used, or a sample of records can be selected.	• Is collected on a routine basis • Is collected by virtually all programs • Is affordable • May be used for longitudinal or panel studies • Allows study of past trends	• May not be complete or accurate • Is limited to few indicators • May not be accessible • May bring up ethical issues of confidentiality
Review administrative reports and documents	• Periodic activity reports • Logistics records • Supply or stock inventory forms • Service delivery guidelines • Supervisory reports • Financial records • Personnel records	• Provides context for program and policy • Documents program history • Tracks trends	• May not be well-maintained • May require extra staff time to assemble or extract from files • Requires permission from manager to access data
Review event logs or other types of logs	• Include information on actions taken, when, to or with whom, by whom, where, how many participated and out comes achieved (i.e., change in program, policy or practice) • Other logs include those for ongoing services provided, media coverage and/or resources generated	• Is easy to complete	• May not be kept up-to-date • May not be representative
Review other documents	• Program records • Correspondence • Official reports	• Provides view of program development and history	• May not be filed centrally • Requires permission from manager to access data

Data Collection Methods: Potential Sources, Advantages and Disadvantages

Data Collection Method	Sources of Data	Advantages	Disadvantages
Conduct interviews with key informants	• Identify knowledgeable individuals to provide information on context and meaning of events • Use interview guides to get information from key informants about the history of the initiative and to identify factors that affected its success or failure	• Provides detailed, "inside" information • Can provide relatively quick assessment of program	• Interviewees may not be well-informed • Interviewees may be biased • Presence of interviewer may influence responses
Conduct surveys*	• Facility surveys • Community surveys • Client follow-up surveys • Consumer or client intercept surveys • Provider surveys • Panel surveys	• Can be tailored to meet specific evaluation needs • Is easy to analyze	• May require external help for sampling, implementation or analysis • May require considerable time to plan and implement • Information may be missed if spontaneous remarks are not recorded
Conduct population surveys*	• General population surveys, nationally representative or local to particular cities or regions	• Is comprehensive • Is useful to monitor change, pre- and post-intervention • Is useful to define program baseline	• Careful sample selection is needed • May not be able to disaggregate data for program sites • May require external help for sampling, implementation and analysis
Conduct exit interviews with clients*	• Client interviews	• Provides direct and immediate client feedback • Is suitable for use with people who are illiterate	• Clients may be reluctant to speak openly • May show "courtesy bias," offering a response thought to be wanted by interviewer
Use mystery clients*	• Mystery client interviews or questionnaires	• May reveal biases or prejudices of the provider	• May be difficult to recruit suitable "mystery clients" in small communities

*Denotes that an illustrative data collection instrument is included in Part II.

Data Collection Methods: Potential Sources, Advantages and Disadvantages

Data Collection Method	Sources of Data	Advantages	Disadvantages
Solicit unstructured feedback from clients	• Informal interviews • Comment and suggestion cards, boards or boxes	• May provide insight into problems not identified by staff	• Can be biased, depending on who voluntarily provides feedback
Conduct site visits	• Clinic flow analysis • Staff interviews • In-depth interviews • Observation • Exit interviews	• Provides on-site evidence of program inputs	• Is time- and labor-intensive • Interpretation of interviews may be subjective and open to bias
Conduct direct observation	• Performance checklists • On-site observation, with the evaluator either observing but not participating, or fully participating	• Provides firsthand assessment of performance in context	• Is time- and labor-intensive • If don't use checklists, interpretation may be subjective • May bring up ethical issues of privacy • May encounter observer bias (i.e., observer only notes what is of personal interest)
Use free-listing	• Respondents' lists of items that they perceive as belonging to the same group	• Facilitates understanding of common themes and concerns	• Language may differ by sub-group
Use pile sorting	• Respondents' characterizations of items	• Facilitates understanding of common themes and concerns	• External reviewers may not understand rationale for piles
Convene focus groups*	• Feedback from focus groups	• Provides in-depth, qualitative information	• Participants may be biased • Requires experienced leader to direct focus group

*Denotes that an illustrative data collection instrument is included in Part II.

Data Collection Methods: Potential Sources, Advantages and Disadvantages			
Data Collection Method	**Sources of Data**	**Advantages**	**Disadvantages**
Create social and census maps	• Maps of facilities and/or that show their relationship to the program sites, the distribution of affiliated programs, barriers to access and other issues	• Provides visual display of local resources	• Different sub-groups (e.g., males, females, adults, youth) may emphasize different map features
Create body maps	• Respondents' body maps show how they understand and interpret anatomy and physiology	• Provides information about level and accuracy of respondents' knowledge of personal anatomy • Reveals local slang and terms used for body parts	• Participants may be shy about speaking of personal topics • Process may be dominated by a few vocal participants
Prepare case studies	• Case studies, describing a series of events or behaviors in detail	• Yields detailed data about program, including perceptions, feelings and social interactions	• Data are unique and not generalizable to other cases
Use population censuses and vital registration systems	• Census and vital registration systems that provide values for most demographic rates, e.g., fertility, mortality, school registration and criminal justice records	• Is suitable for macro-level analysis of major trends	• Coverage of vital records may be incomplete • Censuses typically done only every 10 years

Collecting Data

Many evaluators experienced with collecting data from youth will find that they cannot transfer the methods used with adults to adolescents. The following tips should help you collect information from youth.

Involve community members in the evaluation design and data collection plan.
Involving stakeholders can decrease suspicion of data collectors in the field, reduce opposition to the data collection effort and make parental consent less difficult to obtain. Be prepared to justify to adults why questions they consider "sensitive" need to be asked. Clarify the purpose of data collection and how the information will be used. Be sure to budget enough time for authorization and clearance of survey instruments if that is required.

Carefully define who will participate as respondents.
For data collection, the main issue is defining the lower and upper bound of the age range of youth respondents. You may consider setting a lower bound that relates to the age when parental consent is needed. The upper bound may relate to the trends you want to look at. For example, you may set a higher upper age limit for males in cases when you expect that the male sexual partners of young females are slightly older than the target population you're interested in. To measure impacts that occur 5–10 years after an intervention takes place, you may want to extend the upper age limit of

those interviewed in order to capture longer-term outcomes.

Pay attention to seasonality.
Seasonality affects when youth are available to be interviewed. For example, you may want to collect information during young people's summer vacation from school if you think they are more likely to be at home then. Seasonality may also impact young people's behaviors. For example, in many cultures, marriage ceremonies are timed to take place during an especially "lucky" time of the year; if you survey too close to this time, your data about what percentage of young people are married might be slightly skewed. Similarly, holidays or celebratory festivals often provide young people with opportunities to be sexually active; surveying a month after such an event may yield data that shows higher rates of premarital pregnancy than at other times of the year.

Collect data from places where you know youth will be.
If you plan to conduct a household-based survey, investigate how many youth still live at home. In some communities, large numbers of youth live in boarding houses or on the street, or are in the military. In this case, you might want to sample both from households and from other institutional settings where youth live.

If needed, collect information to help you track adolescents for follow-up.
When conducting certain types of studies, such as panel studies, you will need to interview the same young person at baseline and at follow-up. Locating young people, many of whom are quite mobile, may be difficult. If you are able to collect full names and addresses, that is preferable. If the young person has a nickname, this should also be recorded. Ask young people where they plan to be at the time of the follow-up survey so that you might locate them more easily. Also ask for the names and contact information of one family member and one friend who will always know where they are. In areas where the postal system is reliable, you may want to mail a postcard to youth respondents you plan to follow up on every few months; those that are returned as undeliverable may be youth whom you need to try to locate quickly. Remember, though, that if identifying information is collected from youth, it should be destroyed after the last follow-up data is collected in order to ensure respondents' privacy.

Allow for a high refusal and absentee rate.
After you determine a sample size, you will need to calculate how many youth you will need to approach in order to get the sample size you require. Refusal rates among youth tend to be high—both because parents refuse to give permission, and because young people themselves do not feel comfortable participating in surveys. Absentee rates may also be high. For instance, when conducting a household survey, you may find that young people are rarely at home. For surveys conducted in an institution, such as a school, you should check the average daily absentee rate to estimate how many youth will be present on the day you do the survey.

Budget time for call-backs.
Since absentee rates are high among youth, budgeting time for call-backs is important. You may find that data collectors need to go back to a household or institution two or three times in order to successfully collect information from a young person. You might also have to schedule interviews later in the evening, after youth have returned home from work, or on weekends when they are more likely to have free time.

Build rapport before asking sensitive questions.
Large-scale questionnaires may not be the best way to collect information about highly sensitive information like sexual behavior. It generally takes a while for youth to open up to those collecting data, and it may require either several rounds of data collection or the use of more comfortable qualitative methods to collect honest information about sexual activity.

Ensure privacy.
Ideally, data should not be collected in the presence of other household members. In cases where adults insist on being present during data collection, ask the least sensitive parts of the survey first and hope that the adult leaves the data collector alone by the time more sensitive questions are asked. Some evaluators also ask data collectors to complete an assessment of whether the respondent seemed nervous or uncomfortable during the survey, to check for accuracy of responses.

Provide adequate supervision.
Careful supervision is the key to making sure that data collection procedures are being followed and that problems that arise are resolved in a timely and consistent manner. Having an adequate number of supervisors and field coordinators will help ensure that data is collected in an ethical, systematic and sensitive way. The type of supervision needed will vary by setting and the types of data being collected, but general guidelines are provided above.

Guidelines for Supervising Data Collection

- Choose experienced supervisors who are careful, honest and attentive to detail.
- Periodically review routine data, such as that collected for monitoring, and provide feedback on data quality. To promote data quality, inform those collecting data on how the data are being used to improve program performance.
- Observe data collectors "on the job" at least twice during the course of data collection. You may conduct "spot checks" without informing data collectors when they will be visited.
- On a regular basis, review a sample of completed data collection instruments for completeness and compliance with procedure.
- Be accessible to data collectors and supervisors so that questions and problems can be resolved quickly. Deal with problems as close to the point of data collection as possible so that, if needed, errors can be corrected by returning to the data collection site.

Developing a Workplan for Monitoring and Evaluation

A workplan for M&E will include several kinds of information:

- ➤ tasks involved in carrying out monitoring and evaluation, such as involving stakeholders, assessing the information needs of your project, communicating M&E results and modifying your intervention based on results;
- ➤ timelines for each of these tasks, with a space to check off when each is completed;
- ➤ lists of who is responsible and who will be involved in each stage of monitoring and evaluation; and
- ➤ financial resources needed to complete each task.

Worksheets 7.1 and 7.2 will help you develop your workplan and timetable.

Worksheet 7.1 Evaluation Tasks					
Tasks	**Who Is Responsible for This Task?**	**Who Will Be Involved in This Task?**	**What Equipment Is Needed?**	**What Is the Approximate Budget?**	**What Are Task Start and End Dates?**
1. Specify program objectives					
2. Decide focus and scope of the evaluation					
3. Select indicators					
4. Choose evaluation design					
5. Develop workplan, evaluation team and budget					
6. Collect data					
7. Analyze data					
8. Review and use evaluation results internally					
9. Communicate evaluation results externally					

Worksheet 7.2
Evaluation Timetable

Tasks	Month											
	1	2	3	4	5	6	7	8	9	10	11	12
1. Specify program objectives												
2. Decide focus and scope of the evaluation												
3. Select indicators												
4. Choose evaluation design												
5. Develop workplan, evaluation team and budget												
6. Collect data												
7. Analyze data												
8. Review and use evaluation results internally												
9. Communicate evaluation results externally												

Some evaluations take less than a full year, while others may require two or three years. This worksheet can be adapted to reflect the time frame (in weeks or months) that you estimate will be needed for your M&E effort.

PHOTO: E. Moracco and J. Schelp

8 ANALYZING M&E DATA

CHAPTER AT A GLANCE

- ➤ Details how to process both quantitative and qualitative data
- ➤ Reviews mechanics of data analysis
- ➤ Discusses how to analyze and interpret data to draw conclusions about program design, functioning, outcomes and impact

Processing M&E Data

Processing data refers to the steps needed to organize your data for analysis. This process entails field editing, transcribing, coding, data entry and tabulation and data cleaning, which are each described below. After these five steps, you can move on to data analysis.

Field editing involves reviewing data for completeness and legibility while you are still in the field.
Field editing is the first step in processing qualitative and quantitative data. Field editing involves systematically reviewing field notes; transcripts from focus group discussions, in-depth interviews and observations; and questionnaires.

Data should be reviewed for completeness and legibility while data collectors' memories are still fresh. Reviewing data in the field provides an opportunity to consult the source of the data—facility or a person—in the event that some information is not clear. Field editing also includes the systematic organization of data, recording the date, place and name or other identifier of the informant.

Transcription of qualitative data must be undertaken before data are analyzed.
Transcripts are verbatim records of what was said during a focus group discussion or interview. It is desirable to use a tape recorder to ensure accuracy. If people prefer not to be recorded, have someone take thorough notes. These can then be edited and expanded on while you are still in the field. The transcript will look like a script; it specifies who says what and should also convey notes about gestures or other responses that may not have been recorded on the tape.

Coding refers to a process of organizing and assigning meaning to quantitative and qualitative data.
Data analysis will be simpler if you assign codes to the answers. For example, questions about how much education a young person has completed could have coded responses for each level (e.g., "1=none," "2=primary school," etc.). Most data collection instruments have pre-coded response categories.[1] Use a codebook to

[1] The Comprehensive Youth Survey, Instrument 12, is an example of a pre-coded data collection instrument.

Where Female Newlyweds Say They Seek Treatment for RH Concerns

	1 = Hospital	2 = THC (health complex)	3 = RSDP clinic	4 = Depot holder	5 = Village doctor	6 = *Kabriaj* (traditional healer)	7 = Husband	8 = Sister-in-law	9 = Friends	10 = Mother-in-law	11 = Grandparents
Female discharge			6	5	5	3	6	2		1	
Menstruation problems			3	7	5		5	4			
Safe delivery	6	2	3		1						
Urinary tract infection	1		1	3		1	2	1			
Impotence					4	1		1	2		1
Night emission				1	2	2	1		2		
Tetanus in mother		2	2								
Blood from penis	2				2	1		1	1		
STIs/AIDS	3				1	1					

Note: The higher the number, the more often it was mentioned as a source of treatment.

keep track of how responses to each question have been coded. Add to the codebook as you go along, inserting responses that were not pre-coded by evaluators. All responses, even those handwritten on a structured questionnaire, should be coded and recorded during data entry.

Coding can help organize and interpret descriptive data, such as the answers to open-ended questions about young people's experiences or opinions. After the data are transcribed, each category of response is given a numerical or symbolic code and written in a codebook. When a similar response is found in a subsequent transcript, it is given the same code. For some types of qualitative data collection methods, such as focus groups, transcripts may need to be reduced before they can be coded.

Data will usually be entered into a computer program prior to analysis. When information is collected only from a small number of sites or respondents, it can be tabulated by hand or with a simple spreadsheet program, such as Lotus™ or Excel™. For example, the matrix above was drawn to tabulate results of focus group discussions with female newlyweds in Bangladesh. The left-hand column reflects the reproductive health concerns newlyweds mentioned when asked to free-list their concerns. Each subsequent column reflects answers that were mentioned by at least one respondent in a focus group discussion about where respondents sought

care. The corresponding number reflects the number of focus groups in which this answer was mentioned by at least one participant.[2]

Monitoring data are also often tabulated by hand, using checklists and reports that staff complete, for example, regarding the number of service transactions that took place.

Conversely, most quantitative data is collected from a larger number of respondents and will need to be analyzed with a computer program. Basic spreadsheet programs may be sufficient for smaller data sets. More complex programs, such as *Epi-Info* or *Statistical Package for the Social Sciences (SPSS),* may be needed for larger, more complex data sets. When using computer programs to analyze data, data entry is often time-consuming; for larger evaluations, you may choose to hire outsiders to enter data.

Data cleaning is an essential step. Data cleaning refers to checking for and correcting errors in data entry. Some software packages have built-in systems that check for data entry errors, such as inconsistencies between data items, data omissions and values entered that are out of the range possible. These systems can significantly reduce the amount of time you spend cleaning data.

To check for data entry errors, you should periodically take a sample of data collection instruments and check to see if they are entered correctly. The most rigorous way to eliminate data entry errors is to enter the data twice, then compare the two sets of data item by item. If it is not feasible to do this for all data, than apply this procedure to a sample of cases.

[2] Unpublished analysis by Irit Houvras, Assessment of the Pathfinder Bangladesh Newlywed Strategy, August 1999.

Types of Errors to Be Considered in Data Cleaning

Missing data: Missing data is the result of a respondent declining to answer a question, a data collector failing to ask or record a respondent's answer or a data entry staff member skipping the entry of a response.

Inconsistent data: Within one person's survey, responses are sometimes not consistent. For example, a respondent might say that he had never had sex and then report that he had two sexual partners. The problem should be reconciled by referring to the original questionnaire, if possible. If the respondent's answers are indeed inconsistent, develop a rule about which response to accept.

Out-of-range values: Some data items may be impossible or implausible. For example, "35" is recorded for a 15-year-old female to the question, "How many times have you been pregnant?" Refer to the original survey. If the respondent did give an impossible or implausible answer, you can code the response "no number."

Analyzing M&E Data

Once data are collected and prepared, they can be analyzed. Data analysis will enable you to assess whether and how your program has achieved both program-level and population-level objectives.

In baseline surveys, analysis can reveal:

- participants' characteristics in terms of gender, age, marital status, schooling status, residence and other important attributes; and
- the frequency of specific behaviors and risk and protective factors.

In monitoring and process evaluations, analysis can reveal:

- program quality, coverage and exposure;
- program functions.

In outcome and impact evaluations, analysis can reveal:

- if and how the program achieved its intended results; and
- what portion of the changes in outcome indicators your program can take credit for.

Analysis of data will also enable you to make the following comparisons:

- actual results versus program targets,
- actual progress to projected time frame
- results across program sites, and
- program outcomes versus control or comparison group outcomes.

Analyzing data will enable you to assess whether and how your program has achieved its objectives.

Analyzing Qualitative Data

Some *qualitative data* that you collect will not be coded into numbers and tabulated, but, rather, coded as categories and presented as a narrative or in other forms. You will want to systematically review these data to identify patterns and explore ideas to explain or interpret those patterns. This type of analysis should reflect the original objectives of the program, as well as the evaluation questions posed.

You can present this data in a number of ways:

- *Case studies* are based on transcripts of respondents' narratives. They present one person's interpretation of a program, permitting an in-depth understanding. "Cases" can be individuals, organizations, programs or groups.
- *Process analysis* depicts visually, and with narrative description, a program's processes, or stages of implementation, and how these are linked to outcomes. Process analyses are often presented as flow charts or other graphics, and illustrate how youth programs function and what types of action are required to bring programs about.
- *Causal flow charts* depict sequences of events, revealing how things work or how processes occur by representing actions and events with boxes, circles and arrows. A causal flow chart can be included as part of a process analysis, or can be used to explain how people interpret cause and effect. Another form of causal flow chart that is useful to a youth program is a *decision-tree model,* which graphically outlines the realm of choices and priorities that go into youth's decisions.
- A *taxonomy* is a visual representation or diagram developed by an evaluator to illustrate how respondents relate categories of language and meaning. For example, after collecting data from youth about their reproductive health problems, an evaluator would draw a diagram that illustrates the terms youth use to describe their anatomy and how they understand the link between reproductive health problems and the causes. This taxonomy could be used to assess youth's understanding of reproductive health problems before and after participating in a program, or to compare participants' knowledge with that of non-participants.

Analyzing Quantitative Data

Analysis of *quantitative data* involves further mathematical calculation to produce statistics about the tabulated data. While many want to avoid complicated statistics, much of the analysis done in the typical monitoring and evaluation effort is in fact quite straightforward and involves common sense. This section discusses two types of statistics: *descriptive statistics* and *inferential statistics*.

Calculating descriptive statistics is the first step in data analysis.
Descriptive statistics are used to describe the general characteristics of a set of data. Descriptive statistics include frequencies, counts, averages and percentages. You can use these methods to analyze data from monitoring, process evaluation, outcome evaluation and impact evaluation that have been quantified.

A *frequency* states a univariate (single variable) number of observations or occurrences. For example, when you say that 37 youth of the 242 interviewed have completed the eighth grade, you are stating a frequency. When the frequencies related to a single variable are listed together, this is referred to as a *frequency distribution* (e.g., you may find of the 242 youth interviewed, 37 completed the eighth grade, 148 completed the ninth grade and 57 completed the tenth grade). You can further tabulate data related to more than one variable. For example, you might find that of the 37 youth who completed the eighth grade, 10 are girls and 17 are boys. This is referred to as a *bivariate* or *multivariate* (two or more variables) frequency distribution. Bivariate and multivariate frequencies can be cross-classified and presented in a table. This display of labeled rows and columns is a *cross-tabulation.*

Frequency Distribution of Highest Level of Education Completed by Out-of-School Youth

Grade Completed	Frequency	Percent
Grade 8	15	10.0
Grade 9	40	26.7
Grade 10	65	43.3
Grade 11	20	13.3
Grade 12	10	6.6
Total	**150**	**100.0**

Percentages are calculated by dividing the frequency in one category by the total number of observations, then multiplying by 100.

Descriptive statistics can be used to identify patterns in the data by certain characteristics.
The box above shows an analysis of data collected during a process evaluation conducted at a job training course for young people. It shows both frequency and percent distributions of the highest level of education completed by the 150 youth who attended the training. Of the youth observed (150), the percent that completed Grade 8 is equal to 15 divided by 150 (x 100), or 10 percent. This table tells us that of 150 youth, 120 (80 percent) had completed Grade 10 or less. From this analysis, evaluators found that the training seemed to attract youth who left school after completing Grade 10. This information can be used by program managers to adjust the content of the training so that it best meets the needs of their clients.

By doing further analysis that included a second variable in the table, evaluators learned even more about the youth attending the training. The box on page 136 shows the percent distributions for the highest level of education completed by 150 youth, separated by gender.

Percent of Out-of-School Females and Males by Highest Grade Completed

Grade Completed	Gender		
	Females (percent)	Males (percent)	Total (percent)
Grade 8	12.9	7.1	10.0
Grade 9	31.4	21.4	26.7
Grade 10	42.9	44.3	43.3
Grade 11	10.0	17.1	13.3
Grade 12	2.9	10.0	6.6
Total	**100.0**	**100.0**	**100.0**

Looking at this cross-tabulation, we can see that a higher percentage of girls left school after completing lower grades than did boys. In this case, we have calculated the percentages down the columns, providing information about the distribution for each gender. If we had calculated the percentages across the rows, we would learn for each grade what percent of those who left school were girls versus boys.

To analyze data that have been presented as descriptive statistics, look for patterns in the data that apply to most or all categories of a characteristic being considered, not just one or two. You don't need to observe every item of information; for example, it is unnecessary to state the proportions of males and females falling into each and every educational level—an overall summary of gender differences will usually suffice. Look for dominant patterns or trends by certain characteristics. For example, you might find that program dropout increases as youth get older, or that less-educated youth are more likely to attend a program.

Calculating inferential statistics is the next step in data analysis.
Inferential statistics allow the evaluator to make inferences about the population from which the sample data were drawn, based on probabilities. Inferential statistics are grounded in the concept of probability, or the likelihood of an event occurring. They rely on *statistical significance,* or a way of "giving odds for or against the probability that something happened strictly by chance."[3] Testing for statistical significance helps ensure that differences observed in data, however small or large, were not due to chance.

For example, suppose that descriptive statistics found that the proportion of youth who used condoms the last time they had intercourse was greater in program schools than in control group schools: 45 percent condom use during last intercourse was reported in program schools versus 35 percent in control schools. You should question whether this is a "real" difference, or whether it could be the result of random measurement error.

To answer this question, you could conduct a statistical test to tell you how likely it would be to observe a difference of this size by random chance alone. Suppose that the statistical test indicated that this difference was significant at the 95-percent level of confidence. This would mean that the likelihood of this difference being due to random chance is only 5 out of 100. Thus, you could conclude with a high degree of confidence that condom use in your program schools was higher than in control schools.

Statistics textbooks can provide you with the information you need to conduct such statistical tests. If a member of your staff has received training in statistics, he or she will likely be able to perform basic statistical tests. If needed, you should be able to find help from faculty at local universities. Other methods used to begin the analysis of your

[3] Krause, 1996.

Methods for Analyzing Quantitative Data					
Processing Method	**What You Do to the Data**	**Where the Data Come From**	**How You Get the Information**	**How the Information Is Presented**	**What You Can Do with the Information**
Tabulating	Add items in columns of register or in survey response	Client records, registers or surveys	Take totals and percentages for each item for a given time period	Tables, bar graphs or pie charts	Compare different members of the same category, such as new clients and continuing users, or users of different contraceptives
Cross-tabulating	Choose two data items to see how they are related	Client records, registers or surveys	Break down items in relation to another item	Two-by-two tables in which one item is the independent variable and the other is the dependent variable	Compare different categories of data, such as age of user and method used
Aggregating	Add individual units for overall picture of area	Totals from sites, clinics or providers	Take totals on different times from each unit and add together to get totals for larger area	Tables, bar graphs or pie charts	Compare total situation with program targets
Disaggregating	Break down total situation into units	Summary forms	Take subtotals of items for specific sub-groups of the population	Tables, bar graphs or pie charts	Examine differences between sub-groups based on age, gender or location
Projecting	Forecast how indicators will change over time	Client records, registers or inventory forms	Calculate rates of change in items during a past period, and examine impact of rates over time period in the future	Bar or line graphs	Predict what project outcomes will be if the situation remains unchanged or if rates are changed

Adapted from Wolff et al., 1991.

data include *aggregation, disaggregation* and *projecting,* each of which is explained in the box above.

Once you have determined the appropriate method of analysis, you can begin to consider how your analysis will inform your program at each stage: design, process and outcome/impact.

Analyzing Baseline Data for Program Design

Analysis of baseline data allows us to understand characteristics of our population, identify behaviors and antecedents and determine program coverage and exposure. These issues are all important in understanding whether a program is achieving its population objectives, and in assessing program outcomes and impact. To illustrate, we present data from a baseline survey conducted in Lusaka, Zambia.

FOCUS on Young Adults, at the request of the Lusaka District Health Management Team, the Central Board of Health, and USAID and its partners, conducted a baseline community survey. The survey was designed to serve as a foundation by which to monitor and evaluate the joint implementation of "youth-friendly" clinic services in urban and peri-urban Lusaka. Data were collected on the basis of personal interviews from a total of 2,500 youth who were randomly selected from four treatment and control groups.

The first step in analysis was to calculate descriptive statistics to show the multivariate frequencies of specific behaviors, by other characteristics (see the box below).

The second step was to calculate inferential statistics to determine the antecedents of specific behaviors. We had designed the survey instrument to look at the social influences on the age a young person has sex for the first time. We produced

Distribution of Adolescents by Age, Sex and Illustrative Characteristics

Characteristics		M 10–14		F 10–14		M 15–19		F 15–19		M 20–24		F 20–24	
		#	%	#	%	#	%	#	%	#	%	#	%
Use condom	Yes	3	12.0	6	26.1	94	35.2	64	28.2	156	43.2	104	27.5
	No	23	88.0	17	73.9	173	64.8	163	71.8	205	56.8	274	72.5
	Total	**26**	**100.0**	**23**	**100.0**	**267**	**100.0**	**227**	**100.0**	**361**	**100.0**	**378**	**100.0**
Last sexual partner	Fiancé / husband	2	8.0	1	4.5	11	4.3	88	39.8	93	26.3	207	55.8
	Boyfriend/ girlfriend	14	56.0	16	72.7	220	84.9	128	57.9	239	67.5	148	39.9
	Other	9	36.0	5	22.8	28	10.8	5	2.3	22	6.2	16	4.3
	Total	**25**	**100.0**	**22**	**100.0**	**259**	**100.0**	**221**	**100.0**	**354**	**100.0**	**371**	**100.0**
Age of last sexual partner	14 or less	15	75.0	2	11.1	44	18.6	0	0	3	0.9	1	0.3
	15–19	5	25.0	13	72.2	177	74.7	51	26.0	192	58.4	5	1.4
	20 or greater	0	0	3	16.7	16	6.7	145	74.0	134	40.7	344	98.3
	Total	**20**	**100.0**	**18**	**100.0**	**237**	**100.0**	**196**	**100.0**	**329**	**100.0**	**350**	**100.0**
Can easily buy condom	Agree	78	29.8	71	28.7	302	69.1	286	60.9	313	75.6	313	71.6
	Disagree	32	12.2	20	8.1	81	18.5	66	14.0	75	18.1	62	14.2
	Don't know	152	58.0	156	63.2	54	12.4	118	25.1	26	6.3	62	14.2
	Total	**262**	**100.0**	**247**	**100.0**	**437**	**100.0**	**470**	**100.0**	**414**	**100.0**	**437**	**100.0**

inferential statistics that compared differences in age at first sex. We found that the majority of young people who had had sexual intercourse had also smoked, drank alcohol, used drugs and earned money in the last month. Further, they knew someone their own age who had sex, and spent most of their free time going out with friends. The majority did not attend school, did not live with both parents and did not think that they could talk with their parents. The table below overviews these influences.

[4] See Chapter 3 for a discussion of setting targets.

Analyzing Data Related to Program-Level Objectives

To measure program objectives, data are collected during monitoring and process evaluations to shed light on the design, systems and program functioning and implementation of a program.

Comparing results to the targets you initially set is a way to assess your program's implementation.[4] This is perhaps the most straightforward way of assessing whether your program is being implemented as planned. The box at the top of page 140 provides a comparison of results to targets that have been expressed as indicators.

Inferential Statistics Identifying the Antecedents of Sexual Intercourse in Lusaka, Zambia

Characteristics			% Who Have Had Sex*	No. of Respondents	Odds Ratio**
Individual Characteristics	Ever smoked	Yes	83.1	348	6.8
		No	41.8	1,560	1.0
	Ever drank alcohol	Yes	75.8	807	7.4
		No	29.8	1,100	1.0
	Age group when first drank alcohol	10–15	71.2	379	0.6
		16–24	80.1	401	1.0
	Ever used drugs	Yes	89.8	177	10.7
		No	45.1	1,729	1.0
	Earned money during last month	Something	77.3	428	4.9
		Nothing	41.3	1,474	1.0
Peer Characteristics	Knows someone of same age who has had sex	Yes	70.9	1,261	39.3
		No	5.8	600	1.0
	Spends time with close friend drinking alcohol	Yes	80.1	146	4.6
		No	47.0	1,666	1.0
	Spends time with close friend going to disco/concert	Yes	64.6	181	2.0
		No	48.0	1,631	1.0
	Spends time with close friend going to movies	Yes	60.4	101	1.6
		No	49.0	1,711	1.0
	Spends time with close friend going to parties	Yes	67.4	224	2.3
		No	47.2	1,588	1.0
	Spends time with close friend watching TV/video	Yes	49.1	432	1.0
		No	49.9	1,380	1.0
Social & Family Influences	Attends school	Yes	31.7	1,023	0.2
		No	71.8	794	1.0
	Lives with father or mother	Both	38.2	814	1.0
		Father	50.7	69	1.7
		Mother	50.4	262	1.6
		Other	60.7	763	2.5
	Father likes talking with you	Yes	39.8	699	0.5
		No	54.8	1,209	1.0

* The figures in this column refer to the percentage of young people in each category of the variable or characteristic in question that reported having had sex. As such, they do not add up to 100 percent.

** The odds ratio indicates the likelihood of an event occurring. The higher the number, the more likely something is to have occurred. For example, of youth who answered "Yes" to the question "Father likes talking to you," there is an odds ratio of 0.5 related to them ever having had sex. This means that youth who answered "Yes" to this question were only 50 percent as likely to have had sex than youth answering "No." Odds ratios can illustrate associations between two factors, such as a father talking to a young person being associated with a decreased likelihood of sexual activity.

Comparing Results to Program Targets

Indicator	Target	Data Source	Actual Performance	Percent of Target Achieved	Action to Be Taken
Management plan written, approved and disseminated	Dissemination by end of first program quarter	Checklist	Disseminated by end of first program quarter	Completed as planned	None
Number of advocacy workshops conducted	Six by end of third program quarter	Service statistics	Four by end of third program quarter	67%	Assign higher staff priority to completing workshops
Number of training courses for service providers conducted	22 providers (2 per facility) trained by end of third program quarter	Service statistics	20 providers trained by end of third program quarter	91%	Schedule training for remaining service providers
Number of service visits by young adults	500 in program year 1	Services statistics	600 in program year 1	120%	None

Adapted from Franco et al., 1993.

You can assess program objectives related to quality and coverage by comparing results to the baseline data.[5]

This can help you assess the quality and coverage of your program. Indicators showing quality and coverage might include knowledge and skill levels of program staff, staff service delivery performance, client satisfaction and program dropout rates. The box at the top of page 141 illustrates how data can be compared to baseline data to measure the quality and coverage of a training workshop.

Combine monitoring data and qualitative data to assess program functioning and processes at any stage of the program.

This type of analysis will tell you how your program is functioning and if and how it is achieving its results. To illustrate how qualitative data can be analyzed and presented in youth program evaluations, we draw on an example from Bangladesh.

In June 1999, FOCUS worked with the Bangladesh Rural Advancement Committee (BRAC) and the Rural Service Delivery Program to conduct an assessment of their Adolescent Family Life Education (AFLE) program. This program provided informal primary education and reproductive health education to poor children (ages 11–15) from landless families. It also featured community mobilization and social action, family life education and referral to clinical services. The assessment reviewed sites at different stages of implementation to find out how the program functioned, to determine whether it was reaching its desired outcomes and to make recommendations on how to expand the program to other areas. One of the key challenges of the evaluation was to use a qualitative approach to describe and capture the process of community involvement and social change.

The evaluators, with the collaboration of BRAC Headquarters in Dhaka, used the framework on the bottom of the next page to identify how program objectives would be measured. (See Table on bottom of p. 141.)

[5] See Chapter 3 for a discussion of baseline data.

Evaluation of a Training Workshop				
Indicator	**Target**	**Data Collection Method**	**Actual Performance**	**Action to Be Taken**
Percent of participants who match specified characteristics	90% of participants should match specified characteristics	• Participant survey	50% of participants match specified characteristics	Improve participant recruitment and screening system
Average quality score given by participants	Average quality score given by participants should be greater than 3 on 5-point scale	• Participant questionnaire • Focus group discussion with sample of participants	Average quality score given by participants is 3.7	Use focus group results to improve quality of training
Percent of learning objectives achieved	100% of learning objectives are achieved	• Checklist analyzing workshop content • Participant pre- and posttest measuring knowledge	• All content was covered in workshop • Participant knowledge increases by 10%	Look at areas where participant knowledge is still lacking and modify curricula accordingly

Adapted from Brinkerhoff et al., 1983.

Framework for Assessment of BRAC AFLE Program Objectives				
Program Objectives	**Key Indicators**	**Methods**	**Data Sources**	**Analysis**
• Identify and map social structures • Build trust with the community • Engage community in dialogue about the AFLE program	• Who participates in program and their level of participation • Number of mapping activities conducted and groups identified • Number of group discussions held, number and type of participants, and content and substance of discussion • Trust built • Substantive concerns raised by community • Ideas generated by community • Problems discussed and solutions generated	• Review project documents and monitoring records • Social mapping • Venn diagram • Interviews • Group discussions	• Project documents • MIS • Community leaders • Parents • Management committee	• Descriptive statistics of MIS data • Descriptive accounts from qualitative data • Assessment team discussion of findings and agreement on meanings and implications

Evaluators collected data in each of the program's four districts. First, they listed all 175 schools and 39 *pathogar* (library) sites in a sampling frame, noting which were in the first phase of implementation and which were in the second phase. They then selected a total of two communities with *pathogars* and four communities with current schools, half of which were in Phase I and half in Phase II.

They conducted interviews with other key informants in the community, including program staff (e.g., teachers, program organizers), members of the management committee, community leaders and parents. After transcribing, translating, coding and tabulation, evaluators analyzed the material. Findings are presented in the box below.

Analysis of Findings: Program Objectives of the Social Action Strategy of the BRAC AFLE Program

Identification and mapping of social structures: A review of program documents showed that BRAC staff had previously conducted a survey to identify which eligible children were not currently enrolled in primary school. They also met with elected leaders, religious leaders, government school teachers, community leaders and parents to map adult social groups in the community.

Building trust with the community: Records also showed that BRAC staff sought to build trust with the community before establishing the school by holding informational meetings. Community members were also invited to recommend locations for building the school. Social mapping and Venn diagrams revealed that the way BRAC maintains trust with the community is through a school management committee (5–7 parents, community leaders and the school teacher). In-depth interviews with these community members revealed that the AFLE curriculum was first introduced to the management committee for approval. Most members of the management committee were familiar with the program and could list some of the topics covered in the curriculum.

MIS data confirmed that this committee holds monthly parents' meetings. In most areas, the general topics of the AFLE curriculum had been presented to parents during community forums, which parents reported gave them a positive impression of the school and reduced their resistance to teaching children about reproductive health.

In-depth interviews with BRAC staff revealed that in sites where community leaders or parents were concerned about the AFLE curriculum content, the program organizer and teacher met with those individuals informally to discuss their concerns. A written one-page explanation of the program was used by the program organizers to help them explain the need for the AFLE program to the community.

Interviews also revealed that Moslem religious leaders' advocacy for the program built community trust. One mosque leader said that he had answered parental concerns about the AFLE curriculum. "We made clear to the community that if there was anything against our religion [in the curriculum], we would omit those things, as we do not need them. After going through the curriculum, we found nothing anti-religious; rather, the content was useful and acceptable for the boys and girls. [The BRAC] school is doing very praiseworthy work for our community."

Engaging community in dialogue about the AFLE program: MIS data documented that BRAC facilitated community forums during men's meetings, religious meetings, parents' meetings and newlywed couple meetings. Special forums on the AFLE curriculum, held with students, parents' groups and the school management committee, discussed the efforts to create new social norms supportive of adolescent health. Field interviews with parents confirmed that these meetings took place.

The MIS data indicate the number and types of AFLE sessions held between September 1998 and April 1999. A total of 1,174 sessions were held for students, or 90 percent of those planned. There was also an average of one discussion with parents per school or pathogar, and 90 percent of the schools held meetings. After MIS data were compared by district, it was found that the number of parent meetings held compared to the number planned varied quite a bit, as did the attendance by parents in each site.

	% of Planned Meetings Held	Average No. of Parents/Meeting
Site I:	75%	26 parents
Site II:	97%	23 parents
Site III:	114%	26 parents
Site IV:	51%	15 parents

Comparing Indicators Related to Population Objectives Over Time

Indicator	Data Source	Baseline Value	Follow-up Value	Change in Indicator	Statistical Significance of Change*
Total number of service visits to program facilities by youth	Service statistics	1550	4,000	2,550	(Not applicable)
Proportion of youth who rate program services as "high quality"	Exit interviews, client surveys and focus groups	30%	55%	25%	Significant at 95% level of confidence
Proportion of youth who receive HIV workshop who know the primary ways that HIV is transmitted	Pre- and posttest survey of youth who receive HIV workshop	70%	90%	20%	Not significant at 95% level of confidence
Proportion of youth who use a condom at first intercourse	Population-based survey	50%	60%	10%	Significant at 95% level of confidence
Mean number of sexual partners, last 12 months	Population-based survey	2.4	1.6	0.8	Significant at 95% level of confidence

*The level of statistical significance indicates the probability or likelihood that an observed difference is larger than what would be expected to have occured by chance.

Comparing changes in indicators over time is a way to measure program outcomes.

The following are examples of indicators related to population objectives that could be assessed over time:

- 50-percent increase in the number of youth who use your program's services;
- 30-percent increase in the proportion of youth who rate program services as "high quality";
- 15-percent increase in the proportion of youth who know the primary ways that HIV is transmitted; or
- 10-percent increase in the proportion of youth who use a condom at first intercourse.

The box above illustrates how changes in indicators can be compared over time to assess population objectives.

Comparing evaluation results across sites can alert you to problems, as well as excellence in performance.

This type of analysis will not provide all of the information you need to take necessary corrective action, but it will at least call your attention to sites requiring supervisory attention. In the example in the table below, the data show that Site A is not as far

Analysis of Performance Among Program Sites

Indicator	Data Collection Method	Site A	Site B	Site C	Site D
Percent of target number of peer educators recruited and trained	Service statistics	56%	93%	89%	90%
Percent of peer educators who meet performance criteria	Mystery client	56%	93%	89%	90%
Average number of contacts per peer educator	Service statistics	8	29	11	21
Percent of target audience knowledgeable about how HIV is transmitted	Population survey	43%	70%	51%	59%

Comparing Outcomes with a Control Group				
Indicator	**Treatment Group**	**Control Group**	**Difference (Treatment–Control)**	**Statistical Significance of Difference**
Percent of target audience knowledgeable about how HIV is transmitted	75%	70%	5%	Not significant at 95% level of confidence
Percent of target audience that believe they could obtain a condom if needed	90%	75%	15%	Significant at 95% level of confidence
Percent of young adults that feel comfortable using health services	60%	30%	30%	Significant at 99% level of confidence
Percent of young adults using condom at last intercourse	45%	50%	10%	Not significant at 95% level of confidence

along in program implementation, which may explain the lower levels of knowledge about HIV transmission in the target population. Site B seems to be performing well and producing changes in knowledge about HIV transmission in its target audience. Site C has done well in recruiting and training peer educators, but the number of contacts per peer educator is much lower than at other sites (except Site A). Site D shows good program performance, but the level of knowledge about HIV transmission in the target population is still low.

Comparing changes in indicators with a control or comparison group is another way of measuring impact.[6]
Although control groups are generally used to measure program impact on behavior-related outcomes, they may also be used to answer other types of evaluation questions. For example, if a program objective is to increase the use of reproductive health services among youth, trends in the number of youth who receive services at different health facilities over time can be tracked and compared to service statistics in areas not targeted by your program.

Comparing outcomes with a control group to illustrate impact can be relatively simple.
The box above illustrates an analysis of control group data using a randomized experiment.[7] Data would be collected for each of these indicators, through a survey of the target population.

Comparing outcomes with comparison groups using other study designs can be more complicated.
However, because other study designs do not randomly assign respondents to treatment and control groups, it is necessary to take into account the fact that the groups may differ on factors other than exposure to the intervention. This can be done by applying statistical methods to compensate for possible differences between the groups. The purpose of statistically controlling for differences is to answer the question, "Was the change in indicators for the treatment group significantly larger than that for the control group, after differences between experimental groups on these other factors have been taken into account?"

[6] Control groups are discussed in Chapter 5, along with some of the more commonly used study designs for measuring program impact.
[7] Randomized experiments are discussed in Chapter 5.

It is beyond the scope of this Guide to fully describe statistical procedures. Many statistical textbooks are available to help you select appropriate procedures. Readers are referred in particular to the *Handbook for Family Planning Operations Research Design* (second edition) by Fisher et al. (1991) and *Evaluating Family Planning Programs* by Bertrand et al. (1996), listed in the Reference Shelf (Appendix 3) of this Guide, for details about relevant types of analytic procedures for different study designs. These sources present the analytic procedures in easy-to-understand language. You might also consider hiring a consultant to perform the statistical analysis of your data.

Worksheet 8.1 Compare Actual Performance with a Target or Standard				
Indicator	**Actual Performance**	**Standard or Target for Performance**	**Percent of Standard or Target Achieved**	**Action to Be Taken**

Worksheet 8.2
Analyze Changes in Program Indicators Over Time

Indicator	Baseline Value	Follow-up Value	Change in Indicator	Statistical Significance of Difference

Worksheet 8.3 Compare Performances Among Program Sites						
Indicator	Site A	Site B	Site C	Site D	Site E	Follow-up Action Required

9 USING AND DISSEMINATING M&E RESULTS

CHAPTER AT A GLANCE

- ➤ Reviews reasons to use and disseminate M&E results
- ➤ Describes how to use M&E results to improve your program's interventions
- ➤ Offers tips on how to disseminate results to priority target audiences
- ➤ Presents different formats for dissemination of results

Why Use and Disseminate M&E Results?

There are several reasons to use and disseminate M&E results: to improve program interventions, to strengthen programs institutionally, to advocate for additional resources and "youth-friendly" policies and to contribute to the global understanding of what works. Each of these reasons is explained in detail below.

M&E results help improve your program interventions.
Using M&E results keeps you and your staff in a "learning mode" as you gain understanding about how and why your program is working. M&E results also help you make decisions about the best use of resources. For example, outcome and impact evaluations may provide further insight on certain risk and protective factors, thus shaping your future efforts. As staff use results to reflect on the program's implementation and make necessary improvements, they are more likely to feel supported by the M&E process.

M&E results strengthen your program institutionally.
M&E results can help stakeholders and the community understand what the program is doing, how well it is meeting its objectives and whether there are ways that progress can be improved. Sharing results can help ensure social, financial and political support and help your program establish or strengthen the network of individuals and organizations with similar goals of working with young people. By publicizing positive results, you give public recognition to stakeholders and volunteers who have worked to make the program a success, and you may attract new volunteers.

M&E results can be used to advocate for additional resources and "youth-friendly" policies.
Disseminating M&E results can raise awareness of your program among the general public and help build positive perceptions about young people and youth programs. M&E results often shape donors' decisions about resources in terms of what

and how many to allocate to youth programs. Results can also be used to lobby for policy or legislative changes that relate to youth by pointing out unmet needs or barriers to program success.

If you identify problems early in implementation, you can respond promptly by modifying your program strategy, reassigning staff or shifting financial resources to improve the chances of meeting your program goals and objectives.

M&E results contribute to the global understanding of "what works."
By sharing M&E results, you allow others to learn from your experience. The dissemination of M&E results—both those that show how your program is working and those that find that some strategies are not having the intended impact—contributes to our global understanding of what works and what doesn't work in improving young people's reproductive health. This advances the field by building a body of lessons learned and best practices that can strengthen ARH programs around the world.

Using M&E Results to Improve and Strengthen Your Program

M&E results should be disseminated and used on an ongoing basis, right from the beginning. Your quarterly monitoring and process evaluation reports can be summarized and presented to donors and other stakeholders. Outcome and impact evaluation results should also be used in a timely manner in order to ensure that they have a role in improving and strengthening your program.

M&E results should also lead to decisions about changes in program implementation.
Periodic staff meetings devoted to discussing M&E results can engage staff in collectively making program adjustments. If you identify problems early in implementation, you can respond promptly by modifying your program strategy, reassigning staff or shifting financial resources to improve the chances of meeting your program goals and objectives. If you used a participatory evaluation approach, you should ensure that participants are involved in reviewing results and determining how to use them.

Specifically, M&E results can be used to:

- highlight program strengths and accomplishments,
- improve program management and planning,
- identify weaknesses of program implementation,
- determine demand for service modification or expansion,
- assess quality of care,
- identify future research needs, and
- strengthen funding proposals.

M&E results can help you design new or follow-on activities.
Programs often begin on a small scale in order to test their feasibility. Evaluation results document the strengths, limitations, successes or failures of these initial efforts and allow program planners to make objective decisions about which elements of a program to continue, modify, expand or discontinue. Elements that are not very successful but show promise can be modified for improvement. Successful elements can be expanded by:

- ➤ increasing their scale or scope,
- ➤ changing the administrative structure or staff patterns,
- ➤ expanding the audience and/or targeting new audiences, or
- ➤ spinning off separate programs.

Disseminating M&E Results to Others

Disseminating M&E results to those outside your program is often complex because different audiences will have different information needs. You will have more success disseminating results if you involve major stakeholders, budget adequate resources and develop a dissemination plan before results are finalized.

Determine the audience for M&E results and why you want to share them.
Many different audiences will be interested in evaluation results. Locally, there may be interest among community organizations, school administrators, parent groups, youth, religious organizations, health providers, the media, government officials and social service agencies. At the regional or national level, professional colleagues, reproductive health or adolescent advocacy groups, policymakers and funding agencies may need to learn of your results. nternationally, other program staff, youth advocates, reproductive health specialists, program planners and funding agencies will benefit

Stakeholders include program staff, youth, teachers, parents, school and government officials, community leaders, service providers, and donors.

from the additional knowledge gained about ARH activities.

Think about your reasons for sharing results. Do you want to:

- ➤ increase public awareness about young people's reproductive health status and needs,
- ➤ encourage communities to support youth,
- ➤ improve coordination among agencies working with youth,
- ➤ advocate for policy change,
- ➤ encourage that increased resources be allocated to youth programs, or
- ➤ provide lessons learned for both in-country and international programs?

Share both positive and negative findings.
While every program wants to highlight positive findings, sharing results about what didn't work is also important. Stakeholders also need to understand what is and isn't working, to guide their support toward the most effective youth strategies. Further, most donors appreciate a program's willingness to critically review its work; admitting what hasn't worked well will

Sample Report Outline	
Executive summary	Concisely states the most important and useful findings of the report
Introduction	States the scope of the evaluation (its purpose, audience and key questions)
Background	Explains the setting, target population and basis of the program
Methodology	Describes how the evaluation was carried out
Findings or results	Presents findings about program performance, outcomes and impact
Conclusions	States the evaluator's interpretation of findings
Recommendations	Proposes action, based on conclusions
Lessons learned	Describes implications for similar programs in different settings or for your own program's future activities
Unresolved issues	States what remains to be done or examined and poses unanswered questions
Annexes	Offers additional material that explains evaluation methods, data collection instruments, schedules, persons interviewed, documents reviewed, statistical tables and list of acronyms

Adapted from USAID CDIE, 1997.

improve your credibility and help you solicit funds for changes in program strategy.

Tailoring Dissemination of Results to Different Audiences

Many possible channels exist for presenting evaluation results. For some audiences, one approach may be sufficient (e.g., an all-day retreat with program staff). In other cases, you may want to disseminate results via numerous channels to ensure that the message you are trying to communicate reaches your targeted audience. For example, to reach community members, you might prepare a newspaper story and hold an evening meeting.

In order to plan your dissemination, you must also assess:

- what budget is available,
- the cost of preparing and producing dissemination activities, and
- who is capable of carrying out the activities.

Some activities, such as community-level meetings, are relatively inexpensive. Others, such as publications with broad distribution or national-level seminars, can be quite costly.

Dissemination may be carried out by staff members or may be done in collaboration with outside experts. For example, you may have someone on staff who is able to organize a workshop or conference and someone who is able to prepare a report or publication. If you do not have the capacity to prepare a press release or develop a computerized slide presentation, you may require external assistance.

Common Dissemination Formats

The most commonly used formats are written reports, oral presentations, press releases, fact sheets and slide or computer presentations. While these formats differ in length, detail and the amount of technical information, some common elements are:

- logical organization,
- direct and concise language, and
- use of appropriate illustrations and examples.

A written report combined with visual aids is an effective means of disseminating M&E results.
Written reports can be used to provide an update on the program's progress; document evaluation procedures, findings and recommendations; maintain an internal record of evaluation findings for program staff; and publicize important program information and experiences. To write informative reports that people are likely to read, you should:

- ➤ use clear, simple language in the active voice,
- ➤ be brief and to the point,
- ➤ use attractive layouts, including headings, sub-headings and white space,
- ➤ use boxes, bullets, italics and bold fonts to emphasize important points, and
- ➤ use quotes, anecdotes and case studies to put a human face on the statistics you present.

Visual aids such as maps, tables and charts, graphs, and photographs can be used effectively to summarize information and add "life" to a written report:

- ➤ *Maps* can illustrate areas with high rates of adolescent births, low birth-weight babies, many school dropouts or high youth unemployment. They can also be used to show the program location and the projected impact of activities on the target population.
- ➤ *Tables and charts* are often used to show comparisons—e.g., local statistics in relation to state and national figures—or other information, such as a breakdown of teenage births according to the age of the mother.
- ➤ *Line graphs* can be used to illustrate change over a number of years, such as the number of adolescent births over the past 10 years in a community.
- ➤ *Bar graphs* can also illustrate change over time or changes among sub-groups of the target population.
- ➤ *Photographs* can show your program in action, putting a face on the numbers you are presenting and making readers feel more connected to your project. They can also be used to document community participation in program activities.

Tips for Making an Effective Oral Presentation

Spend time preparing:
- Your presentation should be 20–30 minutes maximum, and cover the main points that your audience needs to know
- Practice your presentation aloud

Anticipate the unexpected:
- Anticipate questions and plan answers
- Plan what you will do if you run out of time or if you finish early

Pay attention to detail and avoid easy mistakes:
- Arrive early and be prepared
- Ensure that audiovisual equipment is available and in good working order
- Avoid last-minute rehearsing or organizing of notes

Give more to the audience than it expects:
- Use visual aids, such as slides or graphics
- Provide appropriate handouts
- Allow time for questions
- Use question-and-answer time to fill in details (if the audience asks)

Based in part on Hanson, Hanson, and Stoddard, 1995.

An *evaluation report* should emphasize only the most important and useful findings, highlighting information that you think will shape the decisions made by staff, donors, policymakers, communities and youth. Keep descriptive information, such as the background of the program, to a minimum, as many readers will be familiar with the program. Include an *executive summary*—i.e., an overview of your main findings. This summary should be written so that it can be distributed independently, for example, to policymakers who may be less likely to read a full report.

Oral presentations are another means of disseminating program results. Oral presentations provide a direct, concise overview of your findings and allow for discussion. You or your staff can give presentations at national meetings, in one-on-one meetings with your board of directors or donors or to community forums. Successful presentations are direct and concise and feature visual aids such as

Preparing Slides and Overheads

The text on slides and overheads should:
- use no more than two different fonts,
- use a font size that is large enough to be seen when projected (at least 24 point),
- separate groupings of ideas with white space,
- have seven words or fewer per line,
- use phrases instead of complete sentences, and
- include no more than six or seven lines per slide.

The graphics on slides and overheads should:
- use bullets effectively;
- use art and graphics that face "no" toward the slide content, not "out" toward the slide edge;
- assure there is enough white space around graphics;
- use charts or graphs to enhance data interpretation; and
- use legends, labels and reference points to enhance charts and graphs.

slides or transparencies. Call attention to the most important points and fill in the details when your audience has a chance to ask questions.

Visual aids help maintain the audience's attention during presentations. Slides, overheads and posters—whether computer-generated or hand-drawn—emphasize important points by presenting information in an abbreviated form. Offer an appealing mix of text (words) and graphics (images).

Press releases can generate media coverage of your findings.
As more people gain access to newspapers, radio, television and the Internet, media coverage of your findings is gaining in importance. Many programs find that the most effective way to reach policymakers is to encourage media coverage of their evaluation results.

A *press release* is a concise statement that presents an overview of your evaluation findings, which you give to the media. The media will usually use the release to develop a story, and it may prompt them to seek additional information about your program's activities.

Fact sheets convey findings in a short, concise format.
Fact sheets are especially effective for advocacy, conveying information to policymakers and others who do not have the time to read longer reports. A fact sheet can also be used as a presentation handout or mailed to program stakeholders. Supply bulleted lists of major findings, keeping the list to under two pages in length.

Tips for Writing a Press Release

- Keep the information simple and clear.
- State the most important information in the first lines.
- Present at most three key findings.
- Limit the release to two double-spaced pages.
- Avoid technical or statistical terms.
- Include the date and information about whom to contact for more information.
- Send to multiple newspapers or radio stations at the same time.
- Send to producers and editors, as well as their staff reporters.

TABLES OF ARH INDICATORS

CHAPTER AT A GLANCE

- ➤ Presents four tables of ARH indicators
- ➤ Features indicators for each phase of a program (program design, program systems development and functioning, program implementation and program intervention outcome)
- ➤ Describes how to use the Indicator Tables

Where Are the Indicators in the Tables From, and How Can I Use Them for My Program?

The tables in this chapter aim to be comprehensive. They include more than 200 indicators from youth programs all over the world, collected during a FOCUS on Young Adults' review of published and unpublished evaluations of youth programs.[1]

The tables offer a wide variety of indicators that must be selected and adapted to your program. In selecting indicators, refer to the Logic Model in Chapter 2 and to your program design. Only measure indicators that you think are relevant and workable for your target community.

What Kinds of Indicators Will I Find in Each of the Four Tables?

The four Indicator Tables offered in this section are based on the different stages of a program that you can measure throughout the course of an M&E effort. These stages (or aspects), introduced in Chapter 3, include:

- ➤ Program Design,
- ➤ Program Systems Development and Functioning,
- ➤ Program Implementation, and
- ➤ Program Intervention Outcome.

Boxes on the next page describe each table in more detail.

[1] The bibliography cites the published results.

Indicator Table I: Program Design Indicators

These indicators are related to standards of quality and efficacy.

Youth programs should be designed based on standards of quality. Standards of quality, while not yet well-tested, are often based on the international experience of youth programs, lessons from the field of family planning and your own intuition and experience. Design indicators will measure whether, and to what degree, quality standards are used to design youth programs.

Examples of design indicators include the following:

- Existence of clearly defined goals and objectives (Yes/No)
- Local stakeholders involved in program planning (Checklist)
- Assessment of needs and preferences of target youth audience for reproductive health services (Yes/No).

Indicator Table II: Program Systems Development and Functioning Indicators

These indicators are related to program objectives and activities.

Program objectives state results in terms of the structure, management or operations of a program; their corresponding activities are related to the development and functioning of your systems. Systems development and functioning indicators measure whether a program's systems are operating and how effectively they have prepared program personnel for implementation. Examples of systems development and functioning indicators include the following:

- Number of peer educators trained to provide youth counseling
- Existence of a clear organizational structure
- Number of partnerships, networks or coalitions established to support the ARH program

Indicator Table III: Program Implementation Indicators

These indicators are related to both program and population objectives and activities.

Both program and population objectives can be fulfilled with the implementation of activities. Implementation indicators measure whether and how many planned activities have been conducted. Examples of implementation indicators include the following:

- Number of youth who seek peer counseling services
- Number and type of involvement by stakeholders in the ARH program
- Number and type of communication products developed for target audience

Indicator Table IV: Program Intervention Outcome Indicators

These indicators are related to population objectives.

Population objectives are measurable statements that state results in terms of the program participant. They measure the changes in outcomes that your program is trying to produce in the target population. Examples of outcome indicators include the following:

- Average age at sexual debut
- Percentage of youth who say they would advocate healthy behaviors among their peers and friends
- Pregnancy rate among female youth during a specified time period
- Incidence rate of STIs for youth during a specified time period

What Other Information Will I Find in the Indicator Tables?

Each table includes the following five columns of information:

Indicators	Calculation*	Data Sources	Notes	Data Collection Instruments

*This column does not appear in Indicator Table I, as no calculation is required.

Below we explain the information that is contained in each column and provide an example.

The first column specifies the *Indicators*.
Indicators are the measures that are suitable to use in monitoring or evaluating your program. Some of the indicators in the tables are accompanied by *symbols.* The following list explains what each symbol means:

- *Ch* reminds you to consider collecting information on the characteristics of the individuals specified in the indicator, such as age and gender of youth or the educational status and residence of family members of youth.
- *G* reminds you to collect information separately based on the gender of youth. Obviously, some indicators refer only to females (e.g., "% of youth who dropped out of school because of pregnancy"). Other indicators are stated in a way that includes both males and females (e.g., "% of youth who ever were pregnant or caused a pregnancy"). For the indicators that have this symbol, however, it is recommended that data be recorded separately for males and females.
- *TD* is a reminder for you to consider adding a time dimension to adapt or specify a given indicator.
- *Cy* reminds you to consider collecting information on the characteristics of youth specified in the indicator, such as age, gender, residence, marital and educational status or group membership. For example, the indicator "% of sexually active youth who have ever used contraception" might be modified as "% of sexually active unmarried girls between 15 and 19 who have ever used contraception."

***Calculation* is specified in the second column.**
The calculation column shows the information needed and the formula for calculating percentages and rates for particular indicators. For example:

This formula might be given to calculate the

$$\frac{\text{No. of youth who have had sex}}{\text{All youth in target population}} \times 100$$ (numerator / denominator)

percentage of youth who are sexually active. If an indicator does not require calculation, "N/A" (not applicable) appears in this column.

***Data Sources* are specified in the third column.**
The Data Sources column suggests different sources from which you may be able to collect information on the indicator, or suggests that you may need to calculate an indicator.

***Notes* are specified in the fourth column.**
The Notes column clarifies or provides additional information about the indicator, or asks you to consider particular issues or topics related to the indicator. This column frequently provides a list of categories or

topics related to one of the terms in the indicator, suggesting that you should consider the topics listed and select those that fit your needs.

Indicators	Calculation	Data Sources	Notes	Data Collection Instruments
% of youth who can identify risk-taking behaviors	No. of youth who have had sex (numerator) / All youth in target population (denominator) x100	Self-reported responses from surveys, interviews with youth	Risk-taking behaviors include: • early age at sexual initiation, • multiple sex partners, and • unprotected sex.	

The last column offers one or more possible *Data Collection Instruments* for a given indicator. Samples of each type of data instrument are provided in Part II of this Guide. This column also indicates if there are specific questions you can use or adapt to get information on a particular indicator. If the indicator is so straightforward that no data collection instrument is necessary, this column will state "N/A."

Indicator Table I: Program Design Indicators

Indicators	Data Sources	Notes	Data Collection Instruments
CROSS-CUTTING DESIGN PROCESS ELEMENTS			
Baseline assessment to identify ARH issues, needs and target audiences	Program log books, program records	Needs of different types of youth should be considered: those not sexually active, those sexually active before marriage and those who are considered high risk. Also, assessment should seek to identify risk and protective factors for youth and look at contexts for risk-taking (e.g., frequent unprotected sex) and health-seeking behaviors (e.g., use of traditional health services).	*For assessment:* Resnick et al. (1997): Findings from the National Longitudinal Study on Adolescent Health Instrument 1A: Program Design Checklist
Political feasibility analysis	Program log books, program records	Analysis includes the following: • Assessment of stakeholder and gatekeeper support • Review of policies and regulations that might limit ARH programs and activities: – Are there overall national/district/local adolescent health policies? – Are these policies supportive to program activities? – Are they strictly enforced? – Do these policies need to be changed in order for the program to begin? – Do new policies need to be developed? • Assessment of collaborative arrangements with other ARH programs and activities and/or between education and service programs	Instrument 1A: Program Design Checklist
Existence of clearly defined mission statement that contributes to the achievement of program goals	Program documents	A mission statement contributes to the achievement of program goals if it: • is consistent with institutional/ local/regional policies and priorities regarding RH, • provides a vision of the future, • defines the program's services and products, • defines the program's target audience/clients, and • is sufficiently clear and detailed in that it provides a meaningful basis for developing operational plans.	Instrument 1A: Program Design Checklist

Indicator Table I: Program Design Indicators

Indicators	Data Sources	Notes	Data Collection Instruments
CROSS-CUTTING DESIGN PROCESS ELEMENTS			
Existence of clearly defined goals and objectives	Program documents	Goals must define the target audience and the ultimate achievement of the program. Objectives must be specific, measurable and time-bound.	Instrument 1A: Program Design Checklist
Intervention goals and strategy based on conceptual model of behavior change	Program documents	Refer to Chapter 2.	Instrument 1A: Program Design Checklist
Local stakeholders involved in program planning	Program log books, records	Stakeholders include the following groups: • Youth • School administrators, teachers and staff • Parents, relatives, caretakers and guardians (referred to as "families or family members" in the rest of the tables) • Community elders/leaders • Policymakers/local government leaders (whatever level is appropriate) • Key social-group representatives (e.g., age-mates in certain cultures, members of religious community, YWCA/ YMCA/Red Crescent, other youth clubs) • Service providers (e.g., nurses, counselors, physicians, social workers)	Instrument 1B: Checklist of Stakeholder Involvement Instrument 2C: Tally Sheet for Stakeholder Involvement Instrument 4: ARH Coalition Questionnaire
Existence of plan to mobilize community	Program documents	Plan includes these activities: • Identify and recruit local stakeholders and community leaders to participate in community activities • Establish a coordinating body (e.g., a committee) to support ARH program activities • Schedule community activities (e.g., community meetings, fairs, festivals) that include ARH messages and information about ARH program activities	Instrument 1A: Program Design Checklist

Indicators	Data Sources	Notes	Data Collection Instruments
CROSS-CUTTING DESIGN PROCESS ELEMENTS			
Assessment of staffing needs and related training requirements	Program documents	Assessment criteria include the following: • History of training for present staff position • Number of existing staff • Staff's competence and confidence with youth • Staff's knowledge in relevant RH subjects • Future training needs and training selection criteria	Instrument 1A: Program Design Checklist
Financial feasibility analysis	Program documents, surveys, budgets	Analysis includes these activities: • Estimating costs • Assessing potential clients' ability and willingness to pay through baseline assessment • Review of financial support sources (e.g., client fees, MOE or local school's operating budget, MOH's budget, NGOs, international and local donors)	Instrument 1A: Program Design Checklist
PROGRAM-SPECIFIC DESIGN PROCESS ELEMENTS			
HEALTH FACILITY PROGRAMS			
Assessment of needs and preferences of target youth audience(s) for RH services	Surveys, exit interviews, interviews with key informants, observations of health facility	Assessment criteria include the following: • Clinic hours and locations • Constellation of services (i.e., what RH and other services and activities are available; for family planning and STI prevention, the range of contraceptive methods available, including condoms) • Services separated for adolescents or combined with other age groups • Characteristics of providers (e.g., gender, age, knowledge, trustworthiness, communication skills) • Type of counseling services (e.g., peer and/or adult, directive or non-directive) • Cost of services	Adapt Instrument 1A: Program Design Checklist for health facilities Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics

Indicator Table I: Program Design Indicators

Indicators	Data Sources	Notes	Data Collection Instruments
PROGRAM-SPECIFIC DESIGN PROCESS ELEMENTS			
HEALTH FACILITY PROGRAMS continued			
Assessment of physical infrastructure for ARH services	Surveys, interviews with key informants, observations of health facility, checklists	Assessment criteria include the following: • Dedicated space or rooms for ARH services • Policies, procedures and space to ensure privacy and confidentiality • Space for audiovisual presentations, counseling and peer education • "Youth-friendly" decor	Adapt Instrument 1A: Program Design Checklist for health facilities Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics
Existence of plan to create "youth-friendly" environment for ARH services	Program documents	Aspects of a "youth-friendly" environment include the following: • Convenient clinic location(s) (e.g., at or near schools, recreation centers or youth centers) • Reachable by affordable transportation • Convenient clinic hours (e.g., before/ after school, evenings, weekends) • Policy and procedures to permit drop-in clients	Adapt Instrument 1A: Program Design Checklist for health facilities Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics
Existence of plan to select providers/staff with "youth-friendly" characteristics	Program documents	"Youth-friendly" characteristics include the following: • Sex • Age (this should not be a primary selection criterion, as an older, confident person may be a better choice than a younger, inexperienced one) • Knowledgeable of youth needs • Friendliness and responsiveness to youth • Trustworthiness • Good communication skills with youth	Adapt Instrument 1A: Program Design Checklist for health facilities Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics Instrument 1E: Checklist of Selection Criteria for Peer Educators
Existence of plan to train providers/staff to serve youth clients	Program documents	Training plan includes the following: • Selection criteria • Appropriate curriculum that is competency-based • Timing/scheduling	Adapt Instrument 1A: Program Design Checklist for health facilities Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics

Indicators	Data Sources	Notes	Data Collection Instruments
PROGRAM-SPECIFIC DESIGN PROCESS ELEMENTS			
HEALTH FACILITY PROGRAMS continued			
Assessment of advisability and level of fees for ARH services	Consumer surveys, exit interviews, interviews with key informants	Youth, clinic managers and service providers should be involved in assessment of fees. Data include the following: • Youth's willingness to pay • Disposable income of youth • Number of youth in target population who have regular income • Whether youth have control over their income	Adapt Instrument 1A: Program Design Checklist for health facilities
Existence of plan to coordinate clinic services with outreach activities	Program documents	Plan includes the following: • Criteria for referrals for inter-/intra-organizational linkages • Calendar of activities/events	N/A
SCHOOL-BASED PROGRAMS			
Mission statement and goals of the program supported by school policies	Institutional records, interviews with school officials and staff	Policy examples include the following: • School RH program is officially endorsed • No eligibility requirements for program activities or services • School supports program for students who are pregnant or parenting	Adapt Instrument 1A: Program Design Checklist for school-based programs
Placement of ARH curriculum in the overall school program, based on review of strategies by key stakeholders	Institutional records, interviews with key stakeholders	Key stakeholders include the following: • School administrators • Teachers • Other school staff • Parents • Community leaders • Students	Adapt Instrument 1A: Program Design Checklist for school-based programs Instrument 1B: Checklist of Stakeholder Involvement
ARH curriculum of adequate duration to meet program goals	Institutional records; interviews with school officials, teachers and staff; observation of curriculum	"Adequate duration" will vary from place to place; 14 hours is considered a minimum.	N/A

Indicator Table I: Program Design Indicators

Indicators	Data Sources	Notes	Data Collection Instruments
PROGRAM-SPECIFIC DESIGN PROCESS ELEMENTS			
MASS MEDIA AND SOCIAL MARKETING PROGRAMS			
Communication strategy design that applies marketing and communication methods	Program documents	Strategy includes the following: • Formative assessment of target audience • Segmentation of target audience • Analysis of behavioral practices (e.g., age of first sex, use of contraceptives, unplanned pregnancies, unsafe abortions) • Analysis of factors influencing youth behavior (e.g., opinions of peers, media images of sexuality, access to contraception) • Development of messages specific to youth and other sub-groups • Pre-testing of messages • Development of plan for contact with media, policymakers and other stakeholders • Use of reinforcing channels of communication (e.g., radio, television, computers, newspapers, magazines, billboards, direct mail, telemarketing systems)	Adapt Instrument 1A: Program Design Checklist for mass media and social marketing programs, using the information from the preceding Notes column

Indicators	Data Sources	Notes	Data Collection Instruments
PROGRAM-SPECIFIC DESIGN PROCESS ELEMENTS			
OUTREACH AND PEER EDUCATION PROGRAMS			
Existence of plan to select youth as peer educators	Program documents	Selection criteria include the following: • Commitment to good reproductive health • Credibility as a role model, especially in regard to behaviors the program advocates • Respect for peers • Ability to hold confidences • Excellent social skills (e.g., ability to interact with both peers and adults, honesty, caring, trustworthiness) • Communication skills (e.g., ability to speak in public and hold public's interest, ability to be understood) • Age, language, geographic location and ethnicity that are similar to target audience	Instrument 1E: Checklist of Selection Criteria for Peer Educators
Existence of plan to train peer educators, adult counselors and program managers to address special needs of youth	Program documents	Training plan includes the following: • Selection criteria • Appropriate curriculum that is competency-based • Timing/scheduling	N/A
Existence of refresher training plan to provide new information and strengthen skills	Program documents	Training plan includes the following: • Selection criteria • Appropriate curriculum that is competency-based • Timing/scheduling	N/A

Indicator Table I: Program Design Indicators

Indicators	Data Sources	Notes	Data Collection Instruments
PROGRAM-SPECIFIC DESIGN PROCESS ELEMENTS			
OUTREACH AND PEER EDUCATION PROGRAMS continued			
Existence of specific objectives and tasks identified for and agreed upon by peer educators	Program documents	Objectives/tasks include the following: • Attending a minimum number of peer education sessions/month (both one-on-one and group sessions) • Attending team meetings • Basic record keeping • Keeping up-to-date on topics covered in sessions • Providing referrals • Mentoring other peer educators • Distributing educational materials and condoms	Adapt Instrument 1A: Program Design Checklist for outreach/peer education programs
Existence of plan for regular supervision and support of peer educators	Program documents	Plan includes the following: • Schedule of visits • Identification of responsible person • Goals of supervision • Means of feedback by supervisors	Adapt Instrument 1A: Program Design Checklist for outreach/peer education programs
Existence of plan to develop educational and presentation materials for use by peer educators	Program documents	Materials include the following: • Posters • Brochures • “Tool kits” with demonstration contraceptives • Body maps	Adapt Instrument 1A: Program Design Checklist for outreach/peer education programs
Existence of plan for presentations/sessions by peer educators that allows for interaction with target audience	Program documents	Interaction activities include the following: • Role playing • Group discussions • Participatory learning for action games	Adapt Instrument 1A: Program Design Checklist for outreach/peer education programs

Indicators	Data Sources	Notes	Data Collection Instruments
PROGRAM-SPECIFIC DESIGN PROCESS ELEMENTS			
YOUTH CENTER PROGRAMS			
Existence of plan to handle turnover of peer educators in program's recruitment, training, scheduling and supervision activities	Program documents	Plan should permit annual or biannual recruitment to accommodate dropouts and peers who "age out" of the program.	Adapt Instrument 1A: Program Design Checklist for outreach/peer education programs
Existence of plan to select youth center staff with "youth-friendly" characteristics	Program documents	"Youth-friendly" characteristics include the following: • Commitment to good reproductive health • Credibility as a role model, especially in regard to behaviors the program advocates • Respect for youth • Ability to hold confidences • Friendliness and responsiveness to youth • Gender that matches target audience	Adapt Instrument 1A: Program Design Checklist for youth center programs Refer to Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics
Assessment of activities and services of interest to both young males and females	Baseline assessment reports	Activities and services include the following: • Health services, such as pregnancy testing and STI prevention, diagnosis and treatment or referrals • Contraceptive/family planning counseling and supply • Nutrition counseling • Social services (e.g., counseling for potential school dropout, job seeking, sexual abuse, family conflicts) • Educational services (e.g., academic tutoring, vocational/job skills) • Recreation	Adapt Instrument 1A: Program Design Checklist for youth center programs
Feasibility analysis of activities and services for young males and females	Review of costs of services, budget, program records	Plan of activities and services must consider program's capacity in terms of staff, facilities or space, timing and costs.	Adapt Instrument 1A: Program Design Checklist for youth center programs

PHOTO:JHU/CCP

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
MANAGEMENT				
Existence of workplan for ARH program activities	N/A	Program documents, MIS	Plan includes the following: • Quantitative or qualitative goals and targets • Strategy for scaling up and/or replication of ARH program, if appropriate • Realistic timeline that accounts for holiday periods, seasonal weather variations that may affect activities, school schedules, etc.	N/A
Achievement of workplan targets	N/A	Program documents, MIS	Program workplan targets include quantitative or qualitative goals to be achieved at different points in program progress. In other words, are the activities being accomplished as planned?	N/A
Adequate financial plan and funding for ARH program activities	N/A	Program documents	Plan includes the following: • Strategy to obtain funds • Strategy to cover recurrent capital costs • System for financial management and accounting • Consideration of long-range sustainability	N/A

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
MANAGEMENT				
Existence of monitoring and evaluation plan	N/A	Program documents	Plan includes these activities: • Assessing program needs for evaluation • Specifying program objectives • Deciding focus and scope of evaluation • Selecting indicators • Choosing a research design • Developing a workplan and budget • Collecting and analyzing data • Communicating evaluation results • Reviewing and using evaluation results	N/A
Existence of clear organizational structure	N/A	Program documents	Structure includes the following criteria: • Program leader is committed to the mission/objectives, ensures that the program is result-driven and that appropriate results are being achieved, and sets an example for the organization • Relationships among organizational units are defined • Roles and responsibilities of staff are defined • Lines of authority are defined	N/A

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
MANAGEMENT				
Adequate staffing in place	N/A	Program documents (especially job descriptions, assessment of current staff and plans for recruitment)	"Staffing" refers to the following: • Number and type of staff positions identified to carry out activities and achieve goals • Existence of skills-based position descriptions showing qualifications, competencies and/or characteristics needed for each position • Existence of recruitment plans • Number of staff positions filled as planned • Assessment of staff competency and performance at periodic intervals	Instrument 6: Inventory of Facilities and Services, Section 8 Instrument 8: Interview Guide for Staff Providing RH Services, Questions 20–27
Access to and use of current information on key operational areas	N/A	Program documents, MIS	Management has access to information on the following: • Staff • Facilities and equipment • Materials and supplies • Commodities and logistics • Finance • Service statistics (This assumes a functioning system of monitoring and evaluation.)	Adapt Instrument 6: Inventory of Facilities and Services by adding management-related items

TD = Time dimension

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
PROGRAM SITES, EQUIPMENT AND SUPPLIES				
Adequate facilities and equipment	N/A	Observation checklists, interviews with program managers and service providers	Assessment includes the following: • No./% of program sites that have types and quantity of equipment sufficient to provide RH services • No./% of program sites where equipment is maintained and functioning, both clinic equipment and audiovisual equipment for presentations	Instrument 6: Inventory of Facilities and Services, Sections 1 and 2
Adequate materials and supplies	N/A	Observation checklists, interviews with program managers and service providers, invoices, MIS	Materials and supplies include items such as pamphlets, flipcharts and videos, as well as clinic supplies such as contraceptives, STI test kits, pregnancy test kits, etc. Assessment looks at the following: • System for monitoring and maintaining materials and supplies • No./% of program sites equipped with needed materials and supplies • System to deliver and distribute supplies to program sites • Percent of staff (e.g., peer educators, supervisors) who have the necessary materials and supplies • Non-burdensome procedures for purchase of supplies (e.g., reasonable number of bids, minimum level of expenditure needing supervisor's authorization)	Instrument 6: Inventory of Facilities and Services, Sections 1 and 3

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
STAFF DEVELOPMENT AND SUPERVISION				
Staff development plan covering pre-service, in-service and refresher training	N/A	Program documents	Plan needs to have specified training objectives to enable monitoring and evaluation.	N/A
Existence of selection criteria for staff training	N/A	Program documents	Criteria cover: • which staff to be trained, • what type of training, and • timing of training.	N/A
Development and testing of training curricula and materials for both training of trainers (TOT) and staff	N/A	Program documents, records, report of pretests	Curricula include: • clinical skills, • interpersonal skills, and • innovative counseling techniques.	N/A
No. of courses that have or cover: • specific learning objectives, • methodology that is interactive, involves role playing and is competency-based, and • relevant RH subjects and issues.	N/A	Review of course curricula or manuals, observations of course presentations	Relevant RH subjects include the following: • Basic RH care • Adolescent physical and sexual growth and development • Predictors of adolescent sexual behavior • Contraceptive methods, including emergency contraception • HIV/AIDS • Sexually transmitted infections • Relationship among risk-taking behaviors (e.g., sexual activity, smoking, drugs, alcohol) • Counseling adolescents • Services for adolescent males • Partner counseling with adolescent patients • Youth culture and language (especially related to sex)	Instrument 1C: Training Course Checklist for ARH Program Staff
No./% of staff trained according to position descriptions	If %: No. of staff trained according to position descriptions x 100 All staff eligible for training	Program records, interviews with trainees	Staff include: • counselors, • nurses, • doctors, • teachers, and • outreach workers.	N/A

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
STAFF DEVELOPMENT AND SUPERVISION				
No./% of trainees who are competent in working with youth	If %: No. of trainees who are competent in working with youth / All trainees x 100	Interviews with trainees, observations of trainees via mystery clients, direct observation of service providers, exit interviews with youth clients	This should be assessed after training has been completed and then reassessed every six months (as resources permit) against the following criteria: • Treats youth with respect and dignity • Applies communication and interpersonal skills • Appears comfortable discussing sexual topics with youth • Is culturally sensitive and respectful of different values, beliefs and practices • Responds appropriately to those who have been coerced or abused	Instrument 7: Observation Guide for Counseling and Clinical Procedures Instrument 8: Interview Guide for Staff Providing RH Services Instrument 9: Guide for Client Exit Interview Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics
No./% of trainees who have mastered relevant subjects	If %: No. of trainees who have mastered relevant subjects / All trainees x 100	Interviews with trainees, observations of trainees via mystery clients, exit interviews with youth clients	This should be assessed after training has been completed and then reassessed every six months (as resources permit) against the following criteria: • Conveys full and accurate information on counseling topic • Counsels appropriate contraceptive methods • Makes appropriate referrals, as needed	Instrument 7: Observation Guide for Counseling and Clinical Procedures Instrument 9: Guide for Client Exit Interview Instrument 10: Questionnaire for Debriefing Mystery Clients
No./% of trainees (i.e., providers) who are competent in their service area	If %: No. of trainees (i.e., providers) who are competent in their service area / All trainees x 100	Interviews with trainees, observations of trainees via mystery clients, exit interviews with youth clients	Assessment requires competency-based criteria that will be specific to each service area.	Instrument 7: Observation Guide for Counseling and Clinical Procedures Instrument 9: Guide for Client Exit Interview Instrument 10: Questionnaire for Debriefing Mystery Clients

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
STAFF DEVELOPMENT AND SUPERVISION				
No./% of trainees who apply new skills to subsequent work	If %: No. of trainees who apply new skills to subsequent work x 100 All trainees	Interviews with trainees, observations of trainees via mystery clients, exit interviews with youth clients		Instrument 7: Observation Guide for Counseling and Clinical Procedures Instrument 9: Guide for Client Exit Interview Instrument 10: Questionnaire for Debriefing Mystery Clients
No./% of trainees who remain active by type of staff	If %: No. of trainees (of a given staff type) who remain active x 100 All trainees (of a given staff type)	Staff records, observation	Efforts should be made to keep youth-dedicated staff in post for duration of program, if they are performing well.	N/A
No./% of staff who receive information to keep up-to-date on relevant subjects	If %: No. of staff who receive information to keep up-to-date on relevant subjects x 100 All staff	Interviews with staff; surveys of staff; documentation such as newsletters, technical updates, directives from Ministry of Health or sponsoring institution		Instrument 8: Interview Guide for Staff Providing RH Services, Questions 24 and 25
No./% of staff who receive supervisory visits [TD]	If %: No. of staff who receive supervisory visits x 100 All staff in need of supervision	Supervisor's records, log books, interviews with supervisors and staff		N/A

[TD] = Time dimension

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
COMMUNITY MOBILIZATION				
No. of local leaders/stakeholders recruited for ARH program	N/A	Program documents	This includes all local leaders (e.g., teachers, health personnel, religious leaders) who will assist the program in whatever capacity.	Instrument 2C: Tally Sheet for Stakeholder Involvement Instrument 4: ARH Coalition Questionnaire Instrument 14: Assessing Coalition Effectiveness Worksheet, Section I
Establishment of a coordinating body (e.g., ARH committee) to support ARH program activities	N/A	ARH committee records, calendar of past and planned activities, interviews with members of coordinating body and other key informants	This refers to the formation of a group of local leaders and other community members to help plan and implement ARH activities at the local level.	Instrument 2C: Tally Sheet for Stakeholder Involvement Instrument 4: ARH Coalition Questionnaire Instrument 14: Assessing Coalition Effectiveness Worksheet, Section II

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
COMMUNITY MOBILIZATION				
No. of partnerships, networks or coalitions established to support the ARH program	N/A	Interviews with key informants (e.g., local leaders), participant observation	Users of this indicator should try to go beyond a simple count of such links to examine whether there is a strategic sharing of resources. For example, if information, materials or messages are shared across organizations in a network, the overall use and effectiveness of the materials should increase. This is done by assessing partnerships against the following criteria: • Organizations volunteer their own funds or sponsor activities • Existing groups expand their mission to include RH for youth • Groups formerly operating separately now collaborate with other groups to further program goals • Collaboration and consensus are the predominant means of decision making • Each group has a specified role • Groups communicate regularly with one another	Adapt Instrument 2C: Tally Sheet for Stakeholder Involvement Instrument 4: ARH Coalition Questionnaire Instrument 14: Assessing Coalition Effectiveness Worksheet
HEALTH FACILITY PROGRAMS				
Facilities conveniently located for youth	N/A	Baseline assessment reports, mapping	"Conveniently located" depends on perceptions and needs of youth, determined through needs assessment. It could mean being near schools, universities, recreation centers and/or affordable transportation. It could also mean a place where youth feel like they have privacy.	Instrument 9: Guide for Client Exit Interview, Questions 53, 54 and 62

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
HEALTH FACILITY PROGRAMS				
Facilities with separate and/or convenient hours for ARH	N/A	Observations of facilities; baseline assessment reports; interviews with managers, service providers and youth	Convenient hours may be before or after school, early evenings and/or weekends.	Instrument 6: Inventory of Facilities and Services, Question 1i Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics Instrument 9: Guide for Client Exit Interview, Questions 50–52
RH service protocols adapted for youth needs	N/A	Program documents, records, review of baseline assessment reports	Protocols should recommend appropriate contraception (e.g., condoms and/or pills as opposed to IUDs); protocols should not require parental or spousal consent in order to provide RH services to youth.	Instrument 6: Inventory of Facilities and Services, Section 5
Private consultation or examination rooms for youth included in facilities	N/A	Observations of facilities, mystery clients, exit interviews, interviews with managers and service providers		Instrument 9: Guide for Client Exit Interview, Questions 23 and 41 Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics
Adequate quality of RH counseling by staff	N/A	Observations of staff by expert, mystery clients, exit interviews	Adequate quality includes the following criteria: • Coverage of essential points in ARH service protocol • Demonstration of appropriate counseling techniques • Development of rapport with target youth audience • Explanation of potentially frightening medical procedures, such as medical exams, both in advance and throughout the procedure • Availability/use of quality IEC materials	Instrument 5A: Index on Quality of Counseling Instrument 7: Observation Guide for Counseling and Clinical Procedures

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
HEALTH FACILITY PROGRAMS				
No./% of program sites with trained health care providers or formal links to nearby providers	If %: No. of program sites with trained health care providers or formal links to nearby providers ___ x 100 All program sites	Observations of program sites, interviews with managers and service providers		Instrument 6: Inventory of Facilities and Services, Section 8 Instrument 8: Interview Guide for Staff Providing RH Services, Questions 22–27
Clinic's existence, location and hours promoted/publicized to potential youth clients	N/A	Surveys, interviews with youth clients and clinic managers, sign board, records of notices, articles, ads on clinic services		Instrument 6: Inventory of Facilities and Services, Question 1h Instrument 9: Guide for Client Exit Interview, Question 21
CURRICULUM DEVELOPMENT (IN SCHOOL-BASED AND PEER EDUCATION PROGRAMS)				
Teaching curriculum and materials developed and tested	N/A	Program records		N/A
Curriculum based on students' needs, interests, sexual experience, ages and cultures	N/A	Baseline assessment reports, observations of curriculum, presentation of curriculum	Curriculum includes the following topics: • Negotiation skills development • Self-esteem • Basic RH care • Adolescent physical and sexual growth and development • Contraceptive methods, including emergency contraception • HIV/AIDS • Sexually transmitted infections • Services for male and female youth	Adapt Instrument 1A: Program Design Checklist for curriculum development

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
CURRICULUM DEVELOPMENT (IN SCHOOL-BASED AND PEER EDUCATION PROGRAMS)				
No. of courses that have or cover the following: • Particular teaching methods to involve participants and personalize the information • Competency-based exercises • Relevant subjects • Activities to address and/or practice key issues and skills • A minimum of 14 hours of instruction • Small-group work with a leader for each group	N/A	Observations/evaluations of course materials and course presentations	This indicator should act as an index of the bulleted points. Some things to consider: • Does participant involvement include such activities as role playing, group discussions and participatory learning for action methods? • Do relevant subjects include reducing sexual behaviors that lead to unintended pregnancy or HIV/STIs, as well as basic RH, risks and methods to avoid risk? • Do key issues include dealing with social influences and pressures via communication, negotiation and assertiveness skills?	Adapt Instrument 1C: Training Course Checklist for ARH Program Staff for curriculum development
MASS MEDIA AND SOCIAL MARKETING PROGRAMS				
No. of communication products that are pre-tested among the intended population	N/A	Program log books	Communication products include radio spots, TV story lines, advertising images and slogans, logos, photonovellas, etc.	Adapt Instrument 2B: Tally Sheet for Communication Products for mass media products
No. of promotional products developed and/or procured	N/A	Program records	Promotional products include key chains, T-shirts, caps, etc.	Adapt Instrument 2B: Tally Sheet for Communication Products for mass media products
No./% of distributors trained in ARH issues and contraceptive use	If %: $\frac{\text{No. of distributors trained in ARH issues and contraceptive use}}{\text{All distributors}} \times 100$	Program records, interviews with or surveys of distributors	Distributors should be sensitive to youth concerns in order to influence sales outlet owners and plan appropriate promotional events.	Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics Instrument 1E: Checklist of Selection Criteria for Peer Educators
No. of contacts with media representatives, policymakers and other stakeholders as ARH program developed	N/A	Program documents, records	This is important as a means of building consensus, engaging potential opponents in constructive dialogue early in the process.	Adapt Instrument 2C: Tally Sheet for Stakeholder Involvement for mass media and social marketing programs

Indicator Table II: Program Systems Development and Functioning Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
OUTREACH AND PEER EDUCATION PROGRAMS				
No. of peer educators recruited and trained	N/A	Program records, log books		N/A
Adequate quality of youth counseling by peer educators	N/A	Observations of peer educators by expert, mystery clients, exit interviews	"Adequate quality" is assessed by the following: • Coverage of essential points in ARH service protocol • Demonstration of appropriate counseling techniques • Coverage of essential information during counseling sessions • Development of rapport with target youth audience • Availability/use of quality IEC materials	Adapt Instrument 5A: Index on Quality of Counseling for peer educators Instrument 7: Observation Guide for Counseling and Clinical Procedures
YOUTH CENTER PROGRAMS				
Youth center conveniently located for youth	N/A	Baseline assessment reports	"Conveniently located" depends on perceptions and needs of youth, determined through needs assessment. It could mean being near to schools, universities, recreation centers and/or affordable transportation. It could also mean a place where youth feel like they have privacy.	Adapt Instrument 9: Guide for Client Exit Interview, Questions 53 and 54, for youth centers
Existence of trained staff, facilities, equipment and supplies for planned youth activities	N/A	Baseline assessment reports, program records, observations of youth center		Adapt Instrument 6: Inventory of Facilities and Services for youth centers
Youth center's existence, location and hours promoted/publicized to potential youth clients	N/A	Surveys, interviews with youth clients and youth center managers, sign board		Instrument 6: Inventory of Facilities and Services, Question 1h Adapt Instrument 9: Guide for Client Exit Interview, Question 21, for youth centers

PHOTO: Harvey Nelson

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
COUNSELING (BY STAFF)				
No. and content of RH counseling sessions held for youth TD	N/A	Service statistics	Content includes: • RH, • HIV/AIDS, • sexually transmitted infections, • contraceptive methods, • strategies to avoid unsafe sex, • negotiation skills and self-esteem, and • referrals for additional services. Sessions can be individual or in groups.	Instrument 2A: Monthly Tally Sheet for Counseling Instrument 3A: Reporting Form for Counseling
No. of youth counseled in RH by staff Ch	N/A	Service statistics	Characteristics include: • age, • sex, • education level, • marital status, • residence, • risk status, and • work status.	Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled
No. and content of RH counseling sessions held for families or family members of youth TD	N/A	Service statistics, observations of counselor by expert, interviews with counselors	"Families" includes parents, relatives, caretakers and guardians, and counseling may be conducted in individual or group family sessions.	Adapt Instrument 2A: Monthly Tally Sheet for Counseling for families of youth
No./% of families or family members of youth counseled in ARH issues by staff TD	If %: No. of families/family members of youth counseled in ARH issues by staff / All families or family members of youth in coverage area x 100	Service statistics		Adapt Instrument 2A: Monthly Tally Sheet for Counseling for families of youth

TD = Time dimension; Ch = Key characteristics

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
COUNSELING (BY STAFF)				
No./% of pregnant young women or parenting youth who have participated in parenting courses/sessions [TD]	If %: No. of youth parents who have participated in parenting courses/sessions x 100 All youth parents in coverage area	Service statistics, self-reported responses from surveys, interviews with youth parents, course records	Content includes: • nutrition, • pre/postnatal care, • immunization, • well baby care, • breastfeeding, • growth monitoring, • hygiene and dental care, and • prevention and treatment of illness (e.g., diarrhea, acute respiratory illness). For courses, check if attendance was by both mother and father, or just mother or just father.	Adapt Instrument 2A: Monthly Tally Sheet for Counseling for youth parents
SKILLS TRAINING (NON-REPRODUCTIVE HEALTH)				
No. and type of skills training courses/sessions held [TD]	N/A	Service statistics	Type of skills training covered includes: • decision making, • negotiation, • communication, • goal setting, and • vocational.	Adapt Instrument 1C: Training Course Checklist for ARH Program Staff for youth
No. of youth participating in skills training [CH] [TD]	N/A	Attendance records, participant data records		Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for youth
No./% of youth who report favorably on the skills training [TD]	If %: No. of youth who report favorably on the skills training x 100 All youth who participate in skills training	Self-reported responses from surveys of youth participants	These youth are those who report that they "like" or "appreciate" a skills training session.	Adapt Instrument 12: Comprehensive Youth Survey, Questions 810, 813, and 814 Adapt Instrument 9: Guide for Client Exit Interview for youth

[TD] = Time dimension; [Ch] = Key characteristics

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
COMMUNITY MOBILIZATION				
No. and type of involvement by local leaders/stakeholders in ARH program	N/A	Program review	Type of involvement includes: • giving a speech, • attending meetings, and • providing resources.	Instrument 2C: Tally Sheet for Stakeholder Involvement Instrument 4: ARH Coalition Questionnaire
No. and type of community activities in support of the ARH program	N/A	Logs of program activities, interviews with personnel conducting activities, observation of selected visits and sessions	Community activities include: • home visits, • community meetings, • fairs, • festivals, and • health days that include YA activities.	Adapt Instrument 2B: Tally Sheet for Communication Products to assess community activities
HEALTH FACILITY PROGRAMS				
No./% of youth who received RH services Ch TD	If %: No. of youth who received RH services / All youth in coverage area x 100	Service statistics		Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled
No./% of youth served by facility who report favorably on key service characteristics Ch TD	If %: No. of youth served by facility who report favorably on key service characteristics / All youth in coverage area served by facility who have received key services x 100	Interviews with or surveys of youth clients		Instrument 1D: Checklist for "Youth-Friendly" Service Characteristics Instrument 9: Guide for Client Exit Interview
No. of youth *first* clinic visits by type of RH service(s) provided Ch TD	N/A	Service statistics	RH services include: • STI screenings and/or treatment, • HIV/AIDS testing, • contraceptive counseling and/or method provision, • nutrition counseling, and • pre/postnatal services.	Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled to assess RH service provision
No. of youth *follow-up* visits by type of RH service(s) provided TD	N/A	Service statistics		Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled to assess RH service provision

TD = Time dimension; Ch = Key characteristics

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
HEALTH FACILITY PROGRAMS (continued)				
No./% of youth referrals by source of referral [TD]	If %: No. of youth referred (by a given source of referral) / All youth in coverage area referred to any given facility x 100	Service statistics	Referral sources include: schools, • private providers, and • peer educators. • These referrals can be made to hospitals, clinics or wherever youth go for services.	Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled to assess referrals
% of youth among all clients who received services [TD]	No. of youth who received services / All clients (regardless of age) who received services x 100	Service statistics	This is for clinics serving both youth and adult populations. If numerator and denominator are the same, then the indicator can measure this at one facility, or you can aggregate service facility statistics for a particular group of facilities or all clinics in a given district.	Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled to assess service provision
No./% of staff who welcome and accommodate youth drop-ins [TD]	If %: No. of staff who accommodate youth drop-ins / All staff who serve youth x 100	Observations of facilities, mystery clients, exit interviews		Instrument 9: Guide for Client Exit Interview, Question 52 Instrument 10: Questionnaire for Debriefing Mystery Clients
SCHOOL-BASED PROGRAMS				
No. of courses held for youth [Ch] [TD]	N/A	Service statistics, participant data records		Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for school-based programs
No. of RH curriculum-related sessions held for families or family members of youth [TD] [Ch]	N/A	Service statistics		Adapt Instrument 2A: Monthly Tally Sheet for Counseling for school-based programs
No. of youth who attended and/or completed RH course [Ch]	N/A	Service statistics	This is measured by tracking attendance and the number of contact hours.	Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for school-based programs
No. of youth referred for RH counseling and/or services from RH courses [Ch]	N/A	Service statistics		Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for school-based programs

[TD] = Time dimension; [Ch] = Key characteristics

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
SCHOOL-BASED PROGRAMS				
No. of families and family members who attended RH sessions [Ch]	N/A	Service statistics	Characteristics include: • age, • gender, • relationship to youth participants, • education level, and • marital status.	Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for school-based programs
MASS MEDIA PROGRAMS				
No. and type of communication products developed for target youth audience(s) [TD]	N/A	Program records	Communication products include: • pamphlets, • posters, • videos, • television spots, • radio spots, and • interpersonal activities (e.g., drama, group presentations).	Instrument 2B: Tally Sheet for Communication Products
No. and type of communication products disseminated or used by target youth audience(s) [TD]	N/A	Log books of radio and TV stations (tallying the number of broadcasts of each spot), routine media monitoring and ratings, broadcast monitoring by staff or volunteers, data from program records on number of posters or brochures distributed, clippings, number of communication events carried out	"Disseminated" refers to the distribution of communication products via electronic, print or other media, as well as the implementation of public relations or interpersonal activities.	Adapt Instrument 2B: Tally Sheet for Communication Products for dissemination of communication products
No. and frequency of communication products used by type of media [TD]	N/A	Log books of various media organizations that have been contacted, clippings	These include such things as ads in newspapers, articles in newspapers and magazines and radio and television spots.	Adapt Instrument 2B: Tally Sheet for Communication Products for mass media programs
No./% of youth in target audience who recall an RH intervention or message [Ch]	If %: No. of youth in target audience who recall an RH intervention or message / All youth in target audience x 100	Self-reported responses from surveys, interviews with target youth audience	"Recall" may be spontaneous or aided; it depends on the data collection method.	Instrument 12: Comprehensive Youth Survey, Questions 908–915 Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for mass media programs

[TD] = Time dimension; [Ch] = Key characteristics

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
MASS MEDIA PROGRAMS (continued)				
No./% of youth in target audience who understand a given message [Ch]	If %: No. of youth in target audience who understand a given message / All youth in target audience x 100	Self-reported responses from surveys, interviews with target youth audience	These are the criteria for "understanding a message": • They correctly state what the message is • They can describe the message in their own words	Instrument 12: Comprehensive Youth Survey, Question 912 Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for mass media programs
No./% of youth in target audience who report favorably about an RH message [Ch]	If %: No. of youth in target audience who report favorably about an RH message / All youth in target audience x 100	Self-reported responses from surveys, interviews with target youth audience	These youth are those who report that they "like" or "appreciate" a communication medium and message.	Instrument 12: Comprehensive Youth Survey, Question 913 Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for mass media programs
For monitoring a telephone hotline or call-in radio show: No. of calls by reason for call No. of callers who cite particular sources of information about the hotline or radio show	N/A	Log books and records from hotline manager	Reasons for calling include seeking information or advice on STIs, pregnancy, abortion, contra-ception, love life, interpersonal skills, etc.	Adapt Instrument 2B: Tally Sheet for Communication Products for mass media programs
No. and type of promotional activities carried out on the RH activities, services and/or contraceptives	N/A	Log books and records	Promotion refers to the advertisement of or publicity for the RH activities, services and/or contraceptives available in the target area.	Adapt Instrument 2B: Tally Sheet for Communication Products for mass media programs
SOCIAL MARKETING PROGRAMS				
No./% of sites stocked with contraceptives and related educational materials that serve youth	If %: No. of youth-serving sites stocked with contraceptives and related educational materials / All sites with a mandate to provide services to youth x 100	Observations of sites, contraceptive records from sites	Examples of such sites include: • pharmacies, • workplaces, • youth centers, • clinics, • schools, • social clubs, • bars, • drama clubs, and • kiosks, or street or market vendors.	Instrument 6: Inventory of Facilities and Services, Sections 1 and 3

[TD] = Time dimension; [Ch] = Key characteristics

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
SOCIAL MARKETING PROGRAMS				
No. and type of contraceptives distributed or sold to youth	N/A	Records from distributors, interviews with or surveys of distributors		Adapt Instrument 6: Inventory of Facilities and Services, Sections 1 and 3, for social marketing programs
No. of youth who receive contraceptives and related educational materials Ch	N/A	Records from distributors, interviews with or surveys of distributors		Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for social marketing programs
OUTREACH AND PEER EDUCATION PROGRAMS				
No. and content of RH sessions held for youth Ch	N/A	Service statistics, interviews with or surveys of peer educators and youth clients, observations of peer educators by expert	This refers to group sessions; content includes: • RH, • HIV/AIDS, • sexually transmitted infections, • sexual violence/exploitation, • negotiation/self-esteem skills and techniques, • contraceptive methods, and • referrals for additional services.	Adapt Instrument 1C: Training Course Checklist for ARH Program Staff for outreach and peer programs
No. and content of ARH sessions held for families or family members of youth	N/A	Service statistics, participant data records, interviews with or surveys of peer educators and youth clients, observations of peer educators by expert	This refers to group sessions.	Adapt Instrument 1C: Training Course Checklist for ARH Program Staff for outreach and peer programs
No. of youth contacted and counseled by peer educators Ch	N/A	Service statistics (kept by peer educators and their supervisors), participant data records	This refers to group sessions.	Instrument 3B: Peer Educators' Reporting Form Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for outreach and peer programs
No. of families or family members of youth counseled in ARH issues by peer educators	N/A	Service statistics (kept by peer educators and their supervisors), participant data records	This refers to group sessions.	Adapt Instrument 2A: Monthly Tally Sheet for Counseling for outreach and peer programs
No. of ARH-related IEC materials distributed by peer educators	N/A	Service statistics	IEC materials include promotional items such as key chains, T-shirts and caps, as well as other traditional IEC materials (e.g., brochures, posters, flipcharts).	Adapt Instrument 2B: Tally Sheet for Communication Products for outreach and peer programs

TD = Time dimension; Ch = Key characteristics

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
OUTREACH AND PEER EDUCATION PROGRAMS (continued)				
No./% of peer educators who distribute or sell contraceptives	If %: No. of peer educators who distribute or sell contraceptives / All functioning peer educators x 100	Service statistics		Instrument 3B: Peer Educators' Reporting Form
No. and type of contraceptives distributed to youth by peer educators	N/A	Service statistics		Instrument 3B: Peer Educators' Reporting Form
No./% of youth contacted and/or counseled by peer educators who are willing to buy or did buy contraceptives from peer educators	If %: No. of youth contacted and/or counseled by peer educators who are willing to buy or did buy contraceptives from peer educators / All youth in coverage area contacted and/or counseled by peer educators x 100	Service statistics	This indicator includes two different issues: intention to buy contraceptives and actually buying contraceptives. It is up to the users of this guide to determine which one they will measure.	Adapt Instrument 2A: Monthly Tally Sheet for Counseling for peer educators
No./% of youth contacted and/or counseled by peer educators who refer friends to peer educators	If %: No. of youth contacted and/or counseled by peer educators who refer friends to peer educators / All youth in coverage area contacted and/or counseled by peer educators x 100	Service statistics, interviews with or surveys of clients		Adapt Instrument 2A: Monthly Tally Sheet for Counseling for peer educators Instrument 3B: Peer Educators' Reporting Form
No. of youth referred for RH counseling and/or services by peer educators Ch	N/A	Service statistics		Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for peer educators
YOUTH CENTER PROGRAMS				
No. and type of youth center activities carried out	N/A	Service statistics	These activities include sporting events, movies/films, plays, music events, etc.	Adapt Instrument 1C: Training Course Checklist for ARH Program Staff for youth centers
No. of youth participating in youth center activities Ch	N/A	Service statistics		Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for youth centers
No. of youth who "hang out" at youth centers Ch	N/A	Service statistics	This can be listening to music, watching a video, playing a game, etc.	Adapt Instrument 2D: Tally Sheet on Number and Characteristics of Youth Counseled for youth centers

TD = Time dimension; Ch = Key characteristics

Indicator Table III: Program Implementation Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
YOUTH CENTER PROGRAMS				
No./% of youth who report favorably on youth center activities	If %: No. of youth who report favorably on youth center activities / All youth who have participated in youth center activities x 100	Self-reported responses from surveys of youth participants	These youth are those who report that they "like" or "appreciate" an activity organized by a youth center.	Adapt Instrument 12: Comprehensive Youth Survey, Questions 1107–1112, for youth centers Adapt Instrument 9: Guide for Client Exit Interview for youth centers
No./% of youth who would recommend the youth center to a friend	If %: No. of youth who would recommend the youth center to a friend / All youth in coverage area who have participated in youth center activities x 100	Self-reported responses from surveys of youth participants		Adapt Instrument 12: Comprehensive Youth Survey, Questions 1107–1112, for youth centers

TD = Time dimension; Ch = Key characteristics

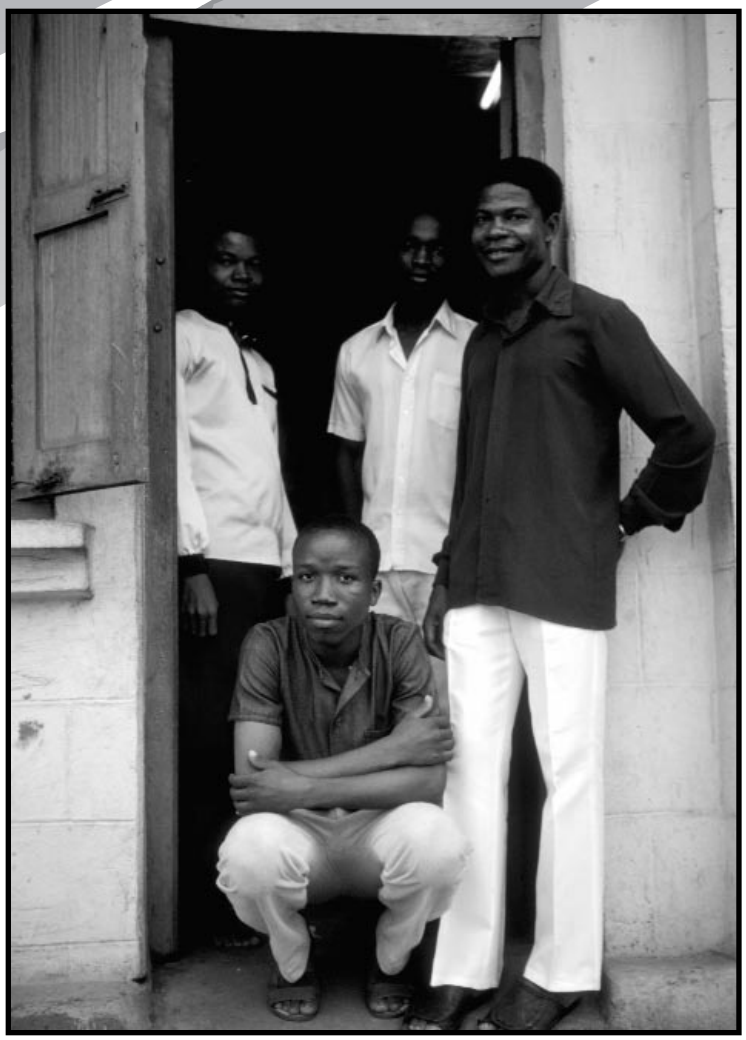


PHOTO:JHU/CCP

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INDIVIDUAL OUTCOME INDICATORS				
KNOWLEDGE				
% of youth who demonstrate knowledge of relevant ARH topics [Ch]	No. of youth who demonstrate knowledge of relevant ARH topics / All youth in target population[1] x 100	Self-reported responses from surveys, interviews with youth, participatory methods	ARH topics include: • basic RH, • contraception, • HIV/AIDS/STIs, • problems associated with pregnancy, and abortion. Topics should also incorporate the relevant services, such as pregnancy testing, STI screening, etc.	Instrument 12: Comprehensive Youth Survey, Modules 2 and 3 Instrument 13: Focus Group Discussion Guide for In-School Adolescents
% of youth who know a source of ARH information and services [Ch]	No. of youth who know a source of ARH information and services / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	ARH sources include: • FLE class instructors, • school nurses, • health clinics, • peer educators, and • youth centers.	Instrument 12: Comprehensive Youth Survey, Modules 2 and 3 Instrument 13: Focus Group Discussion Guide for In-School Adolescents
% of youth who can identify risk-taking behaviors [Ch]	No. of youth who can identify risk-taking behaviors / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	Risk-taking behaviors include: • early age at sexual initiation, • multiple sexual partners, • unprotected sex, and • using drugs/alcohol.	Instrument 12: Comprehensive Youth Survey, Modules 2 and 3
% of youth who can articulate options available to avoid risky behaviors	No. of youth who can articulate options available to them to avoid risky behaviors / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	Options include the following: • Seeking RH services • Staying abstinent • Using contraceptives • Being monogamous • Avoiding drugs and alcohol • Referring peers to RH activities/ services/contraceptives	Instrument 12: Comprehensive Youth Survey, Modules 2, 3 and 4

[1] For calculating these indicators, the target population needs to be defined by incorporating the exact chatchment area as well as specific types of youth at risk.

[Ch] = Key characteristics; [G] = Gender; [TD] = Time dimension; [Cy] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INDIVIDUAL OUTCOME INDICATORS				
ATTITUDES, BELIEFS AND VALUES				
% of youth who have particular attitudes and/or beliefs about key health-related behaviors, influences and issues	No. of youth who have a particular attitude and/or belief about a key health-related behavior, influence or issue / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	Attitudes about the following are relevant: • Relationships, monogamy, serial monogamy, multiple partners and/or marriage • Age at first intercourse • Abstinence and/or sexual activity • Premarital sex • Contraception (e.g., use, acceptability, risks and benefits of modern and traditional methods) • Desired number of children • Getting pregnant within first year of marriage • Exchange of money or goods for sex • STIs and HIV/AIDS • People with HIV/AIDS	Instrument 12: Comprehensive Youth Survey, Module 4
% of youth who have discussed attitudes on key health-related behaviors, influences and issues during an RH intervention	No. of youth who have discussed attitudes on key health-related behaviors, influences and issues during an RH intervention / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	The discussion should have taken place within the past three months and should have been longer than 15 minutes. The discussion can take place between youth and their parents, health care providers, teachers, counselors, peer educators and/or grandparents.	Instrument 12: Comprehensive Youth Survey, Questions 450–452
INTENTIONS				
% of youth who believe that the ideal age of marriage for males is below the average male age for marriage in the country	No. of youth who believe the ideal male age of marriage is below the average male age for marriage in the country / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	The average age of marriage for males will vary, depending on the location. This information can be collected from vital statistics or from a recent national demographic survey.	Instrument 12: Comprehensive Youth Survey, Module 4
% of youth who believe the ideal age of marriage for females is below the average female age for marriage in the country	No. of youth who believe the ideal female age of marriage is below the average female age for marriage in the country / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	The average age of marriage for females will vary, depending on the location. This information can be collected from vital statistics or from a recent national demographic survey.	Instrument 12: Comprehensive Youth Survey, Module 4

CH = Key characteristics; G = Gender; TD = Time dimension; CY = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INDIVIDUAL OUTCOME INDICATORS				
INTENTIONS continued				
% of youth who expect to marry at an early age	No. of youth who expect to marry at an early age x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods	An "early" age is typically considered an age that is below the average age of marriage in a country, region or district.	Instrument 12: Comprehensive Youth Survey, Module 4
% of youth who intend to have sex before marriage	No. of youth who intend to have sex before marriage x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 4
% of youth who have particular intentions about key health-related behaviors	No. of youth who have particular intentions about key health-related behaviors x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods	These include an intention to: • consult an RH provider for information, advice and/or services, • delay onset of sexual activity, • use a contraceptive within a fixed period of time, and/or • use condoms to protect against STIs/HIV.	Instrument 12: Comprehensive Youth Survey, Module 4
% of youth who have discussed their intentions about key health-related behaviors [G]	No. of youth who have discussed their intentions about key health-related behaviors x 100 All youth in target population	Self-reported responses from surveys, interviews with youth	These are the criteria for the discussion: • Occurred within the past three months • Lasted at least 15 minutes • Took place with youth and their parents, health care providers, teachers, counselors, peer educators and/or grandparents	Instrument 12: Comprehensive Youth Survey, Questions 450–452
SELF-EFFICACY				
% of youth who feel responsible for their own welfare and well-being	No. of youth who feel responsible for their own welfare and well-being x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 4
% of youth who feel responsible for their actions toward others	No. of youth who feel responsible for their actions toward others x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 4
% of youth who believe they could refuse sex if they didn't want it	No. of youth who believe they could refuse sex if they didn't want it x 100 All youth in target population	Self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 7

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INDIVIDUAL OUTCOME INDICATORS				
SELF-EFFICACY continued				
% of youth who believe they could get their partners to use contraception/condoms	No. of youth who believe they could get their partners to use contraception/condoms x 100 All youth in target population	Self-reported responses from surveys, interviews with youth	This includes having the confidence to be able to negotiate with partner about using condoms and/or other contraceptives.	Instrument 12: Comprehensive Youth Survey, Modules 4, 6 and 7
% of youth who believe they could seek RH information and services if they need them [G]	No. of youth who believe they could seek RH information and services if they need them x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods	This includes buying contraceptives from peer educators who sell them, visiting health services, attending FLE courses, etc.	Instrument 12: Comprehensive Youth Survey, Modules 4 and 7
% of youth who believe they could advocate particular "healthy" behaviors among their peers, friends and partners	No. of youth who believe they could advocate particular "healthy" behaviors among their peers, friends and partners x 100 All youth in target population	Self-reported responses from surveys, interviews with youth	Healthy behavior includes: • seeking RH services, • staying abstinent, • using contraceptives, • being monogamous, • avoiding drugs and alcohol, and • referring peers to RH, activities/services/contraceptives.	Instrument 12: Comprehensive Youth Survey, Modules 4 and 7
SKILLS				
% of youth who acquired proficiency in a particular skill during the course of a life skills or ARH intervention [G]	No. of youth who acquired proficiency in a particular skill during the course of a life skills or ARH intervention x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods	Skills include: • self-awareness, • empathy, • interpersonal relationships, • decision making, • critical thinking, • problem solving, • negotiation, • coping with stress, • coping with emotions, • communication, • goal setting, and • self-efficacy to get contraceptives and/or treatment.	N/A

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [Cy] = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INDIVIDUAL OUTCOME INDICATORS				
SKILLS continued				
% of youth who feel comfortable discussing RH issues with adults, health providers or peer educators	No. of youth who feel comfortable discussing RH issues with adults, health providers or peer educators / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	RH issues include: • contraceptive methods, • STIs/HIV/AIDS, • pregnancy, • abortion, and • menstruation/wet dreams.	Instrument 12: Comprehensive Youth Survey, Modules 5, 7 and 8
BEHAVIORS				
SEXUAL ACTIVITY				
% of youth who have ever had sexual intercourse	No. of youth who have ever had sexual intercourse / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	"Sexual intercourse" refers to the insertion of the penis into the vagina.	Instrument 12: Comprehensive Youth Survey, Module 6
Age at first intercourse by key characteristics of youth	N/A	Self-reported responses from surveys, interviews with youth, participatory methods	Key characteristics include the following: • Whether they have had any courses in ARH/sex education • Whether they have ever seen an RH provider • Whether they have participated in any youth program • Education level • Current age • Sex	Instrument 12: Comprehensive Youth Survey, Module 6
% of youth who have had sex within a specified time period [TD]	No. of youth who have had sex within a specified time period / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	Although many adult surveys use a standard reporting time period of one to three months (to reduce reporting errors), a longer time period, perhaps one year, may be necessary for adolescents.	Instrument 12: Comprehensive Youth Survey, Module 6
No. of times youth have had sex within a specified time period [TD]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods	Although many adult surveys use a standard reporting time period of one to three months (to reduce reporting errors), a longer time period, perhaps one year, may be necessary for adolescents.	Instrument 12: Comprehensive Youth Survey, Module 6

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [Cy] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
SEXUAL ACTIVITY continued				
No. of sexual partners ever	This indicator can be calculated as a mean: Total no. of lifetime sexual partners Total no. of sexually active youth	Self-reported responses from surveys, interviews with youth, participatory methods	• Differentiate between sex that was coerced and sex that was voluntary or consensual. • This indicator may be calculated as the mean number of partners or as a percentage (for example, percentage of youth who have had more than three partners).	Instrument 12: Comprehensive Youth Survey, Module 6
No. of sequential sexual partners within a specified time period TD	N/A	Self-reported responses from surveys, interviews with youth, participatory methods	• This indicator is used to examine patterns of serial monogamy (when one relationship begins after another has ended). • This indicator may be calculated as the mean number of partners or as a percentage (for example, percentage of youth who have had more than three partners). • Although many adult surveys use a standard reporting time period of one to three months (to reduce reporting errors), a longer time period, perhaps one year, may be necessary for adolescents.	Instrument 12: Comprehensive Youth Survey, Module 6
No. of sexual partners within a specified time period TD	N/A	Self-reported responses from surveys, interviews with youth, participatory methods	• This indicator measures the number of sexual partners an individual has during a specified period of time. • This indicator may be calculated as the mean number of partners or as a percentage (for example, percentage of youth who have had more than three partners). • Although many adult surveys use a standard reporting time period of one to three months (to reduce reporting errors), a longer time period, perhaps one year, may be necessary for adolescents.	N/A

CH = Key characteristics; G = Gender; TD = Time dimension; CY = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
SEX WITH SAME-SEX PARTNERS				
% of young males who have had sexual contact with another male	No. of young males who have had sexual contact with another male / All young males in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
% of young males who have had anal intercourse with another male	No. of young males who have had anal intercourse with another male / All young males in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
% of young males who have ever used a condom for anal intercourse with another male	No. of young males who have ever used a condom for anal intercourse with another male / All young males in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
% of young males who used a condom at last anal intercourse with another male	No. of young males who used a condom at last anal intercourse with another male / All young males in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
No. of same-sex partners [TD]	N/A	Self-reported responses from surveys		Instrument 12: Comprehensive Youth Survey, Module 6
SEXUAL ABUSE, COERCION AND EXCHANGE				
No./% of youth who report having been victims of sexual abuse [CH] [TD]	If %: No. of youth who report having been victims of sexual abuse / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	Sexual abuse includes rape or sex between family members. This applies to both males and females.	Instrument 12: Comprehensive Youth Survey, Module 6
No./% of youth who have ever been forced to have sex [CH]	If %: No. of youth who have ever been forced to have sex / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	This applies to both males and females. Variations of this indicator include the following: • Ever feeling pressured to have sex • Ever being tricked into having sex • Ever having unwanted sex	Instrument 12: Comprehensive Youth Survey, Module 6
No./% of youth who have ever had sex for money or other form of exchange [CH]	If %: No. of youth who have ever had sex for money or other form of exchange / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	This applies to both males and females.	Instrument 12: Comprehensive Youth Survey, Module 6

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
SEXUAL ABUSE, COERCION AND EXCHANGE continued				
No./% of youth who have ever paid money or other form of exchange for sex [CH]	If %: No. of youth who have ever paid money or other form of exchange for sex / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	This applies to both males and females.	Instrument 12: Comprehensive Youth Survey, Module 6
CONTRACEPTION AND CONDOM USE[2]				
% of sexually active youth who have ever used modern contraception [CY] [G]	No. of sexually active youth who have ever used modern contraception / All youth in target population who have had sexual intercourse x 100	Self-reported responses from surveys, interviews with youth, participatory methods	Modern contraception includes: • condoms, • oral contraceptives, • injectables, • foam/jelly, • Norplant, and • spermicide.	Instrument 12: Comprehensive Youth Survey, Module 6
% of sexually active youth who used contraception at first intercourse [CY] [G]	No. of sexually active youth who used contraception at first intercourse / All youth in target population who have had sexual intercourse x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
% of sexually active youth who used contraception before first pregnancy [CY] [G]	No. of sexually active youth who used contraception before first pregnancy / All youth in target population who have been pregnant x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
% of sexually active youth who used contraception at last intercourse [CY] [G]	No. of sexually active youth who used contraception at last intercourse / All youth in target population who have had sexual intercourse x 100	Self-reported responses from surveys, interviews with youth, participatory methods	This indicator can also be calculated with respect to contraceptive method. For example, one may examine the percentage of youth that used condoms at last intercourse.	Instrument 12: Comprehensive Youth Survey, Module 6
% of sexually active youth who are currently using contraception, by method [CY] [G]	No. of sexually active youth who are currently using contraception (for a given method) / All youth in target population who have had sexual intercourse x 100	Self-reported responses from surveys, interviews with youth, participatory methods	This indicator can also be computed by type of method used (e.g., condoms, birth control pills).	Instrument 12: Comprehensive Youth Survey, Module 6

[2] Contraception and condoms are used to prevent unwanted pregnancy, sexually transmitted infections and HIV/AIDS.

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
CONTRACEPTION AND CONDOM USE continued				
% of sexually active youth who carry a condom [Cy] [G]	No. of sexually active youth who carry a condom / All youth in target population who have had sexual intercourse x 100	Self-reported responses from surveys, interviews with youth, participatory methods	Although it is ideal that all youth carry condoms, regardless of sexual experience, programs will need to decide whether it is realistic to use all youth as the denominator for this indicator.	Instrument 12: Comprehensive Youth Survey, Module 6
Source of supply by method for sexually active youth who used a contraceptive/condom at last intercourse or are currently using a contraceptive/condom	N/A	Self-reported responses from surveys, interviews with youth, participatory methods	Sources of supply include health centers, schools, peer educators, social marketing programs, pharmacies, friends, families, etc.	Instrument 12: Comprehensive Youth Survey, Module 6
PREGNANCY				
% of youth who were ever pregnant or caused a pregnancy [Cy] [G]	**Females:** No. of young females who were ever pregnant / All young females in target population x 100 **Males:** No. of young males who have ever caused a pregnancy / All young males in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
Age at first pregnancy for young females	N/A	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
No. of times young females have ever been pregnant	N/A	Self-reported responses from surveys, interviews with young females		Instrument 12: Comprehensive Youth Survey, Module 6
% of youth who have had or caused an unintended pregnancy [Cy] [G]	**Females:** No. of young females who have had an unintended pregnancy / All young females in target population x 100 **Males:** No. of young males who have caused an unintended pregnancy / All young males in target population x 100	Self-reported responses from surveys, interviews with youth	• Depending on the analysis needs of the program, the denominator can also be expressed either as all youth who have had sexual intercourse or all youth who have been pregnant/caused a pregnancy. • "Unintended pregnancy" includes pregnancies that were wanted but mistimed and pregnancies that were unwanted altogether.	Instrument 12: Comprehensive Youth Survey, Module 6

[Ch] = Key characteristics; [G] = Gender; [TD] = Time dimension; [Cy] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
PREGNANCY continued				
% of births to young females that were wanted	No. of births to young females that were wanted / Total number of births to young females in target population x 100	Self-reported responses from surveys, interviews with youth	This indicator may also be calculated for young males who have caused a birth.	
% of young females who avoid repeat pregnancy [CY]	No. of young females who avoid repeat pregnancy / All young females in target population who have been pregnant x 100	Self-reported responses from surveys, interviews with youth	This indicator may also be calculated for young males who have caused a birth.	Instrument 12: Comprehensive Youth Survey, Module 6
% of young females who practice(d) a specified level of pregnancy-related care [CY]	No. of young females who practice(d) a specified level of pregnancy-related care / All young females in target population who have been pregnant x 100	Self-reported responses from surveys, interviews with youth, participatory methods	Pregnancy-related care refers to: • prenatal care, • delivery assistance, • postnatal care, and/or • abortion and post-abortion care.	Instrument 12: Comprehensive Youth Survey, Module 6
% of young females who have ever had an induced abortion [CY]	No. of young females who have ever had an induced abortion / All young females in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	• Depending on the cultural location, it may be difficult to get information on the type of abortion (spontaneous or induced) and whether legal or clandestine. However, if it is possible, all these factors provide valuable information. • This indicator may be calculated for young females who have been pregnant. • This indicator may be calculated for young males who have caused a pregnancy.	Instrument 12: Comprehensive Youth Survey, Module 6
Pregnancy rate among young females during a specified time period [TD]	No. of pregnancies among young females during a specified time period / All young females in target population x 1,000	Ministry of Health records, clinic surveillance data, DHS surveys	These rates can be calculated by age groups, for example, 10–14, 15–19 and 20–24 year olds.	N/A
No. of induced abortions or abortion rate among young females during a specified time period [TD]	If rate: No. of induced abortions among young females during a specified time period / All young females in target population x 1,000	Ministry of Health records, clinic surveillance data, DHS surveys	This information may be impossible to get in some country settings. However, a survey may be possible if one has not already been conducted on this issue.	N/A
Fertility rate among young females during a specified time period [TD]	No. of births to young females during a specified time period / All young females in target population x 1,000	Ministry of Health records, clinic surveillance data, DHS surveys	These rates can be calculated by age groups, for example, 10–14, 15–19 and 20–24 year olds.	N/A

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
PREGNANCY continued				
% of young females who dropped out of school because of pregnancy [Cy]	No. of young females who dropped out of school because of pregnancy x 100 All young females in target population who attend school	Self-reported responses from surveys, interviews with youth, school records, Ministry of Education records	This indicator may also be calculated for young males who have caused a birth and dropped out in order to support their child, or for other reasons.	Instrument 12: Comprehensive Youth Survey, Module 1
Of young females who dropped out of school due to pregnancy, % who returned or intend to return to school	No. of young females who returned or intend to return to school x 100 All young females who dropped out of school because of pregnancy	Self-reported responses from surveys, interviews with youth, school records	Need to specify whether youth mothers completed school (and at which level) or the amount of time the youth mother spent at school.	Instrument 12: Comprehensive Youth Survey, Module 1
SEXUALLY TRANSMITTED INFECTIONS				
% of youth who report specific STI symptoms [Cy] [TD]	No. of youth who report specific STI symptoms x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods, clinic records	This indicator may be calculated as % of youth who have ever had any STI symptoms OR as % of youth who had any STI symptoms in the last year.	Instrument 12: Comprehensive Youth Survey, Module 3
No. of youth who seek treatment for STIs [Cy]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods, clinic records	This indicator may also be calculated by source of treatment, for example, by type of health sector.	Instrument 12: Comprehensive Youth Survey, Module 3
% of youth who were ever diagnosed with an STI [Cy]	No. of youth who were ever diagnosed with an STI x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods, clinic records	A diagnosis of an STI could be either though a laboratory test or a syndromic approach.	Instrument 12: Comprehensive Youth Survey, Module 3
No. of reported cases or incidence rate of STIs among youth during a specified period [G] [TD]	If incidence rate: No. of youth who were diagnosed with an STI during a specified time period x 1,000 All youth in target population	Ministry of Health records, clinic surveillance data, DHS surveys		N/A
Prevalence rate of STIs among youth during a specified period [G] [TD]	No. of youth who have an existing STI plus those who were diagnosed with an STI during a specified time period x 1,000 All youth in target population	Ministry of Health records, clinic surveillance data, DHS surveys		N/A
Of those who were ever diagnosed with an STI, % of youth who received treatment	No. of youth who were ever diagnosed with an STI who received treatment x 100 All youth in target population who were ever diagnosed with an STI	Self-reported responses from surveys, interviews with youth, clinic records	If possible, collect information on the type of treatment.	Instrument 12: Comprehensive Youth Survey, Module 3

[Ch] = Key characteristics; [G] = Gender; [TD] = Time dimension; [Cy] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
SEXUALLY TRANSMITTED INFECTIONS continued				
Of those who were ever diagnosed with an STI, no. of times youth had an STI in the last year [TD]	N/A	Clinic records, follow-up studies	Although self-reported responses may be the most accessible source of data for this indicator, the most reliable source of data is *retrospective* clinical studies of cases of infection.	Instrument 12: Comprehensive Youth Survey, Module 3
Of those who were ever diagnosed with an STI, % of youth who avoid repeat infection [CY] [TD]	No. of youth who were ever diagnosed with an STI who avoid repeat infection / All youth in target population who were ever diagnosed with an STI x 100	Clinic records, follow-up studies	This indicator may be difficult to measure. The characteristics of youth may be difficult to determine from lab reports, and repeat infections may not be possible to confirm. Although self-reported responses may be the most accessible source of data for this indicator, the most reliable source of data is *prospective* follow-up studies of cases of infection.	Instrument 12: Comprehensive Youth Survey, Module 3
DRUG AND ALCOHOL USE				
% of youth who have ever smoked a cigarette [G]	No. of youth who have ever smoked a cigarette / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 10
Age when first tried smoking cigarettes [G]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 10
% of youth who are current smokers [G] [CY] [TD]	No. of youth who are current smokers / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	"Current smokers" refers to persons who have had at least one cigarette in the past week.	Instrument 12: Comprehensive Youth Survey, Module 10
Of those who are current smokers, no. of cigarettes smoked/month [G] [TD]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 10
% of youth who have ever used a substance to make them "high" [G] [CY]	No. of youth who have ever used a substance to make them "high" / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	"Substances" include: • marijuana, • drugs (e.g., amphetamines, barbiturates, opium, hallucinogens), and • glue or gas (or other inhalants).	Instrument 12: Comprehensive Youth Survey, Module 10

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
DRUG AND ALCOHOL USE continued				
No. of times/month youth have used drugs [G] [TD]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 10
% of youth who have ever used drugs to enhance a sexual experience [G] [Cy]	No. of youth who have ever used drugs to enhance a sexual experience / All youth in target population who have ever had sexual intercourse x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 10
% of youth who have ever drunk alcohol [G]	No. of youth who have ever drunk alcohol / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 10
Age when first tried drinking alcohol [G]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 10
% of youth who have been drunk in the past month [G] [TD]	No. of youth who have been drunk in the past month / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	"Drunk" is defined as consuming three to five alcoholic drinks within a three-hour period. This indicator can also be calculated using only youth who have ever had an alcoholic drink in the denominator.	Instrument 12: Comprehensive Youth Survey, Module 10
% of youth who consume five or more drinks in a row [G] [Cy]	No. of youth who consume five or more drinks in a row / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	This indicator can also be calculated using only youth who have ever had an alcoholic drink in the denominator.	Instrument 12: Comprehensive Youth Survey, Module 10
EMOTIONS/BEHAVIORS THAT MAY LEAD TO OTHER RISKY BEHAVIORS				
% of youth who think they might run away from home [G]	No. of youth who think they might run away from home / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	"Running away" is defined as leaving home for an extended period of time without giving prior notification. Usually, the person who runs away is trying to escape from some problem at home.	Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who have run away from home [G]	No. of youth who have run away from home / All youth in target population x 100	Self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 8

[Ch] = Key characteristics; [G] = Gender; [TD] = Time dimension; [Cy] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
EMOTIONS/BEHAVIORS THAT MAY LEAD TO OTHER RISKY BEHAVIORS continued				
% of youth who currently feel depressed [G] [TD]	No. of youth who currently feel depressed / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	Depression is a psychological disorder that typically includes a feeling of sadness, inactivity, lack of motivation to carry out daily activities, a difficulty in thinking and concentrating, and a feeling of rejection.	Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who have ever felt depressed [G]	No. of youth who have ever felt depressed / All youth in target population x 100	Self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who feel stress [G] [TD]	No. of youth who feel stress / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	Stress is a factor that makes a person feel mental or physical tension. Some causes of stress include a death of a family member, receiving low grades in school, having a lot of work with little time to complete it, etc.	Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who currently feel lonely [G] [TD]	No. of youth who currently feel lonely / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	Loneliness is a feeling of sadness from a lack of companionship, and feeling apart from friends and/or family.	Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who have ever felt lonely [G]	No. of youth who have ever felt lonely / All youth in target population x 100	Self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who have contemplated suicide [G] [TD]	No. of youth who have contemplated suicide / All youth in target population x 100	Self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 8
MARRIAGE			These indicators are useful for two reasons: • An objective of some programs is to encourage youth to stay in school and postpone marriage. • Marital status and age at marriage help define the characteristics of the target population.	
% of youth who have ever been married or have cohabited [CY]	No. of youth who have ever been married or have cohabited / All youth in target population x 100	Vital statistics, self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 1

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
BEHAVIORS				
MARRIAGE continued				
Of those who have ever married or cohabited, average age at marriage or co-habitation	N/A	Vital statistics, self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 1
% of youth who are currently married or cohabiting [Cy]	No. of youth who are currently married or cohabiting / All youth in target population x 100	Vital statistics, self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 1
PARENTING				
% of infants or children of youth parents who received the recommended immunizations on schedule [Cy]	No. of infants or children of youth parents who received the recommended immunizations on schedule / All infants or children of youth parents in target population x 100	Self-reported responses from surveys, interviews with youth parents, immunization records	Certain variables must be considered for this indicator: • Adherence to immunization schedule • Receiving immunizations in correct order	Refer to DHS Version 4.0 (September 1998)–Section 4b
% of youth mothers who breastfeed(fed) their infants [Cy]	No. of youth mothers who breast-feed(fed) their infants / All youth mothers in target population x 100	Self-reported responses from surveys, interviews with youth mothers		Refer to DHS Version 4.0 (September 1998)–Section 4a
% of youth fathers who are actively involved in parenting their children	No. of youth fathers who are actively involved in parenting their children / All youth fathers in target population x 100	Self-reported responses from surveys, interviews with youth parents	This includes: • feeding, • bathing, • disciplining, or • educating.	N/A
% of youth fathers who contribute to the financial support of their children	No. of youth fathers who contribute to the financial support of their children / All youth fathers in target population x 100	Self-reported responses from surveys, interviews with youth parents	Financial support may be comple-mented with other material goods and support.	N/A
RELATIONSHIP OUTCOME INDICATORS				
PEER/PARTNER CHARACTERISTICS				
Communication and Information from Peers				
% of youth who have spoken with their peers about sex	No. of youth who have spoken with their peers about sex / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 4

[Ch] = Key characteristics; [G] = Gender; [TD] = Time dimension; [Cy] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
RELATIONSHIP OUTCOME INDICATORS				
Communication and Information from Peers continued				
% of youth who prefer to get information about sex from their peers	No. of youth who prefer to get information about sex from their peers x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 4
Perceptions of Peer Attitudes/Behaviors				
% of youth who perceive that their peers are having premarital sex	No. of youth who perceive that their peers are having premarital sex x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who perceive that their peers think it is wrong to have premarital sex	No. of youth who perceive that their peers think it is wrong to have premarital sex x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who perceive that their sexually active peers are using contraception	No. of youth who perceive that their sexually active peers are using contraception x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who perceive that their peers use alcohol	No. of youth who perceive that their peers use alcohol x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who perceive that their peers use drugs	No. of youth who perceive that their peers use drugs x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who perceive that their peers smoke cigarettes	No. of youth who perceive that their peers smoke cigarettes x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who perceive that their peers visit sex workers	No. of youth who perceive that their peers visit sex workers x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5

Ch = Key characteristics; G = Gender; TD = Time dimension; Cy = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
RELATIONSHIP OUTCOME INDICATORS				
PEER/PARTNER CHARACTERISTICS continued				
Partner Relations[3]				
% of youth who currently have a partner	No. of youth who currently have a partner / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	A "partner" is anyone with whom a youth is romantically and/or sexually involved.	Instrument 12: Comprehensive Youth Survey, Module 1
% of youth who feel committed to their partners	No. of youth who feel committed to their partners / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 1
% of youth who had a partner at age 15 or younger	No. of youth who had a partner at age 15 or younger / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 1
% of female youth whose current partner is at least five years older	No. of female youth whose current partner is at least five years older / All female youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
% of youth who communicate with current partner about particular concerns/topics	No. of youth who communicate with current partner about particular concerns/topics / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	Concerns might include: • RH issues, • sex, and/or • contraception.	Instrument 12: Comprehensive Youth Survey, Module 6
% of youth who have felt pressured by their current partners to have sex	No. of youth who have felt pressured by their current partners to have sex / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
% of youth whose last sex was unwanted	No. of youth whose last sex was unwanted / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 6
% of youth who think it is okay to pressure their partners for sex in some circumstances	No. of youth who think it is okay to pressure their partners for sex in some circumstances / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 4

[3] Many of these indicators are characteristics of partner relations, but an intervention could aim to try to improve partner relations among the target population.

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
RELATIONSHIP OUTCOME INDICATORS				
FAMILY				
Family Dynamics				
% of youth who feel comfortable talking with their parent(s)	No. of youth who feel comfortable talking with their parent(s) / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	A "parent" could also be a grandparent, a guardian or anyone that youth consider to be their primary caretaker(s).	Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who say they feel supported by their parents	No. of youth who say they feel supported by their parents / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who believe they currently have a harmonious relationship with their parents	No. of youth who believe they currently have a harmonious relationship with their parents / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	A "harmonious relationship" refers to the absence of conflict or the ability to resolve conflict.	Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who have ever discussed sexual matters with either parent	No. of youth who have ever discussed sexual matters with either parent / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 5
% of youth who have responsibilities in their households	No. of youth who have responsibilities in their households / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods	Responsibilities include: • taking care of siblings, • cooking, and • fetching water and/or firewood.	Instrument 12: Comprehensive Youth Survey, Module 8
% of families who think youth should conceive immediately after marriage	No. of families who think youth should conceive immediately after marriage / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods	"Families" include parents, adult relatives, caretakers and guardians of youth.	Instrument 15: Parents of Youth Questionnaire
% of families who think their children should have only one or two children	No. of families who think their children should have only one or two children / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who feel comfortable talking with their children [TD]	No. of families who feel comfortable talking with their children / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who want their children to complete secondary school	No. of families who want their children to complete secondary school / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
RELATIONSHIP OUTCOME INDICATORS				
FAMILY continued				
Family Attitudes, Beliefs and Values				
% of families who have ever discussed sexual matters with their adolescent children	No. of families who have ever discussed sexual matters with their adolescent children / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who feel comfortable discussing sexual matters with their adolescent children	No. of families who feel comfortable discussing sexual matters with their adolescent children / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who believe they have enough knowledge to discuss sexual matters with their children	No. of families who believe they have enough knowledge to discuss sexual matters with their children / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who have discussed their values about premarital sex with their adolescent children	No. of families who have discussed their values about premarital sex with their adolescent children / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who are aware of youth organizations in their communities	No. of families who are aware of youth organizations in their communities / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who would support their children's participation in youth organizations	No. of families who would support their children's participation in youth organizations / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods	Youth organizations include both secular and church-related organizations.	Instrument 15: Parents of Youth Questionnaire
% of families who are aware of organizations that provide ARH information to youth	No. of families who are aware of organizations that provide ARH information to youth / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who would allow their adolescent children to access ARH information from an organization	No. of families who would allow their adolescent children to access ARH information from an organization / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
RELATIONSHIP OUTCOME INDICATORS				
FAMILY continued				
Family Attitudes, Beliefs and Values continued				
% of families who think ARH information should be provided in schools	No. of families who think ARH information should be provided in schools / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who are aware of health facilities that provide ARH services to youth	No. of families who are aware of health facilities that provide ARH services to youth / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods	ARH services include: • pregnancy testing, • contraceptive counseling and distribution, • STI screening and treatment, • HIV/AIDS testing, • abortion/postabortion care, and • prenatal/postnatal care.	Instrument 15: Parents of Youth Questionnaire
% of families who would allow their children to seek RH services from a health facility	No. of families who would allow their children to seek RH services from a health facility / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
% of families who think condoms and other contraceptives should be available to youth	No. of families who think condoms and other contraceptives should be available to youth / All families in target population x 100	Self-reported responses from surveys, interviews with families, participatory methods		Instrument 15: Parents of Youth Questionnaire
INSTITUTIONAL OUTCOME INDICATORS				
RELIGIOSITY				
% of youth who say they have religious and/or spiritual beliefs	No. of youth who say they have religious and/or spiritual beliefs / All youth in target population x 100	Self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who believe that religion shapes their attitudes about sexual behavior	No. of youth who believe that religion shapes their attitudes about sexual behavior / All youth in target population x 100	Self-reported responses from surveys, interviews with youth		Instrument 12: Comprehensive Youth Survey, Module 8

CH = Key characteristics; G = Gender; TD = Time dimension; CY = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INSTITUTIONAL OUTCOME INDICATORS				
RELIGIOSITY continued				
% of youth who belong to a religious group	No. of youth who belong to a religious group / All youth in target population x 100	Self-reported responses from surveys, interviews with youth	A "religious group" includes: • a youth group at a church, mosque, temple or synagogue, and • a religious group at or outside of school.	Instrument 12: Comprehensive Youth Survey, Modules 1 and 8
EDUCATION/SCHOOLING				
% of youth who are literate G Cy	No. of youth who are literate / All youth in target population x 100	Ministry of Education records, surveys		Instrument 12: Comprehensive Youth Survey, Module 1
% of youth who attend school G Cy	No. of youth who attend school / All youth in target population x 100	Ministry of Education records, surveys	Schooling includes both academic and vocational training.	Instrument 12: Comprehensive Youth Survey, Module 1
No. of school dropouts or drop-out rates for youth in a period of time Cy G TD	If rate: No. of youth school dropouts in a period of time / No. of youth in target population that attended school in a period of time x 100	Ministry of Education records, school records		N/A
Level of educational attainment for youth G	N/A	Ministry of Education records, surveys	Educational attainment includes both academic and vocational training, and age of educational attainment should be specified. Level of attainment refers to the last year of completed studies.	Instrument 12: Comprehensive Youth Survey, Module 1
% of youth who believe that academic achievement is important	No. of youth who believe that academic achievement is important / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 1
% of youth who want to complete secondary school	No. of youth who want to complete secondary school / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 1
% of youth who want to attend a college/university	No. of youth who want to attend a college/university / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 1
% of youth who have ever felt unsafe in school	No. of youth who have ever felt unsafe in school / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 1

CH = Key characteristics; G = Gender; TD = Time dimension; Cy = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INSTITUTIONAL OUTCOME INDICATORS				
EDUCATION/SCHOOLING continued				
% of youth who are involved in extracurricular activities at school	No. of youth who are involved in extracurricular activities at school _____ x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 1
CONNECTION TO ADULTS/COMMUNITY INSTITUTIONS				
Degree of community support for ARH programs	N/A	Program records	This indicator will be assessed by developing a policy environment score based on the following criteria: • Involvement or support of local leaders/stakeholders in ARH program • Existence of local coordinating body for ARH program activities • Partnerships, networks and coalitions that work to support ARH program activities • Community activities that support ARH program • Advocacy activities in support of ARH program • Local funding for ARH program • Level of stigma associated with ARH issues • Level of opposition to ARH program	Instrument 4: ARH Coalition Questionnaire Instrument 5B: Policy Environment Score: Adolescents Instrument 14: Assessing Coalition Effectiveness Worksheet
No. of key stakeholders involved in ARH programs	N/A	Program records	The relevant issue may not be the number but rather that the most influential key group members are involved. Depending on the context of the program, a count or a proportion of the key representatives should be conducted. For example, if the program's context is schools, then the number of school representatives involved should be obtained.	Instrument 2C: Tally Sheet for Stakeholder Involvement Instrument 4: ARH Coalition Questionnaire

CH = Key characteristics; G = Gender; TD = Time dimension; CY = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INSTITUTIONAL OUTCOME INDICATORS				
CONNECTION TO ADULTS/COMMUNITY INSTITUTIONS continued				
% of youth who are involved in leisure activities for youth outside of school	No. of youth who are involved in leisure activities for youth outside of school x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods	Leisure activities for youth include: • sports, • drama, • music, and • crafts or art.	Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who receive ARH information from organizations outside of school	No. of youth who receive ARH information from organizations outside of school x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 11
% of youth who have sought counseling services from youth organizations	No. of youth who have sought counseling services from youth organizations x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 11
% of youth who have relationships with adults through community or neighborhood institutions	No. of youth who have relationships with adults through community or neighborhood institutions x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods	For example, the adult relationship with youth could be a mentorship or a friendship.	Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who feel supported by other adults in the community	No. of youth who feel supported by other adults in the community x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 8
% of youth who believe they could consult with an adult if they had a problem	No. of youth who believe they could consult with an adult if they had a problem x 100 All youth in target population	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Modules 4 and 7
INSTITUTIONAL INFRASTRUCTURE				
No. of service organizations for victims of sexual abuse in community	N/A	Community surveys, interviews with key informants, checklists		Instrument 2E: Institutional Infrastructure Tally Sheet
No. of youth organizations in community	N/A	Community surveys, interviews with key informants and youth, checklists		Instrument 2E: Institutional Infrastructure Tally Sheet
No. of schools that provide ARH information	N/A	Community surveys, interviews with key informants and youth, checklists		Instrument 2E: Institutional Infrastructure Tally Sheet

KC = Key characteristics; G = Gender; TD = Time dimension; CY = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
INSTITUTIONAL OUTCOME INDICATORS				
INSTITUTIONAL INFRASTRUCTURE continued				
No. of youth organizations that provide ARH information	N/A	Community surveys, interviews with key informants and youth, checklists		Instrument 2E: Institutional Infrastructure Tally Sheet
No. and type of economic opportunities for youth in community	N/A	Community surveys, interviews with key informants and youth, checklists	Economic opportunities can be in an informal or a formal sector of work.	Instrument 2E: Institutional Infrastructure Tally Sheet
No. of entertainment venues targeting youth in community	N/A	Community surveys, interviews with key informants and youth, checklists	Entertainment venues can include: • video houses, • recreation centers, • youth centers, • cinemas, or • theaters.	Instrument 2E: Institutional Infrastructure Tally Sheet
No. of venues that serve alcohol to youth in community	N/A	Community surveys, interviews with key informants and youth, checklists		Instrument 2E: Institutional Infrastructure Tally Sheet
COMMUNITY OUTCOME INDICATORS				
EXPOSURE TO MEDIA				
% of youth who have ever seen a pornographic film, magazine or other form of media	No. of youth who have ever seen a pornographic film, magazine or other form of media / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 9
No. of exposures to pornographic film, magazine or other form of media [TD]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 9
% of youth who have seen violent films, comics or books	No. of youth who have seen violent films, comics or books / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 9
No. of exposures to violent films, comics or books [TD]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 9

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
COMMUNITY OUTCOME INDICATORS				
EXPOSURE TO MEDIA continued				
% of youth who have seen an educational video/film or magazine on an RH issue	No. of youth who have seen an educational video/film or magazine on an RH issue / All youth in target population x 100	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Survey, Module 9
No. of exposures to an educational video/film or magazine on an RH issue [TD]	N/A	Self-reported responses from surveys, interviews with youth, participatory methods		Instrument 12: Comprehensive Youth Survey, Module 9
POLICY				
Degree of political support for ARH policies and programs	N/A	Surveys, program records, interviews with key informants	This indicator covers support by government (at the appropriate level and including policymakers and functionaries), political parties and religious organizations, as well as public opinion and media campaigns.	Instrument 5B: Policy Environment Score: Adolescents
Existence of adequate resources directed to ARH programs	N/A	Surveys, program records, interviews with key informants	Criteria include the following:[4] • Adequacy of funding from government sources • Adequacy of funding from donor sources • Adequacy of staffing for program activities and/or services • Adequacy of activity or service points and staff/providers to ensure reasonable access by most youth in target population • Allocation of resources by explicit priority guidelines	Adapt Instrument 1A: Program Design Checklist to assess resources for ARH programs Instrument 5B: Policy Environment Score: Adolescents, Section V
Existence of minimum legal age for marriage	N/A	Surveys, interviews with key informants, checklists		Instrument 5B: Policy Environment Score: Adolescents
Enforcement of legal age for marriage	N/A	Surveys, interviews with key informants, checklists		Instrument 5B: Policy Environment Score: Adolescents

[4] The POLICY Project.1998. *Program Design and Evaluation Guidelines.*

[CH] = Key characteristics; [G] = Gender; [TD] = Time dimension; [CY] = Characteristics of youth

Indicator Table IV: Program Intervention Outcome Indicators

Indicators	Calculations	Data Sources	Notes	Data Collection Instruments
COMMUNITY OUTCOME INDICATORS				
POLICY continued				
Existence of legal abortion	N/A	Surveys, interviews with key informants, check-lists		Instrument 5B: Policy Environment Score: Adolescents
Legality of contraceptive sales to youth	N/A	Surveys, interviews with key informants, check-lists		Instrument 5B: Policy Environment Score: Adolescents
Legality of condom sales to youth	N/A	Surveys, interviews with key informants, check-lists		Instrument 5B: Policy Environment Score: Adolescents
Existence of rape laws	N/A	Surveys, interviews with key informants, check-lists		Instrument 5B: Policy Environment Score: Adolescents
Enforcement of rape laws	N/A	Surveys, program records, interviews with key informants, police records		Instrument 5B: Policy Environment Score: Adolescents

CH = Key characteristics; G = Gender; TD = Time dimension; CY = Characteristics of youth

GLOSSARY

Access The extent to which services are available at a cost and effort that is acceptable for those who need them.

Adolescence The transition between childhood and adulthood, usually defined as including those ages 10 to 19.

Advocacy The act or process of supporting a cause or proposal.

Aggregate (adj.) Taking all units as a whole.

Aggregate (v.) To collect or gather into a mass or whole.

Analysis Identification of and explanations for patterns of information collected to provide answers to research questions being studied.

Antecedent A factor that precedes and influences (both positively and negatively) how a person makes decisions and behaves.

Assessment A systematic process of gathering information, analyzing it and then making a judgment. In this Guide, assessment is used to mean a process that takes place before a program begins.

Average (n.) The sum of all, divided by the total number of cases. Also called the *arithmetic mean*.

Baseline information Facts and figures collected before an intervention begins.

Baseline survey A structured way of collecting factual information from multiple respondents about the state of a population before an intervention begins.

Bias To influence, distort or prejudice.

Biased sample A sample that is not representative of the population from which it is taken.

Body mapping A data collection method; participants draw maps of the female and male bodies, focusing on the details of the reproductive system and how it functions.

Calculate To determine by mathematical process.

Case study A detailed analysis and description of an event, program, situation, condition or organization in the context of its environment.

Catchment area The geographic territory of a service, program or facility.

Categorical data Data that are not expressed in terms of numbers but as belonging to one category or another (e.g., male or female).

Census mapping A data collection method used to gather information about a geographic area, such as household data on the number of youth, gender of family members, education and literacy levels, employment and resource ownership.

Checklist A list that enumerates key features of a setting or process that users can check off during observations.

Client An individual who receives services or participates in a program.

Cluster sampling A method for approximating a random sample by aggregating individuals in larger units of observation that occur naturally (e.g., classrooms, schools, health facilities, youth groups), taking a random sample of the larger units (clusters) and then observing all or a sample of the individuals within the sampled clusters.

Coalition An alliance of parties or persons for joint action.

Community An interacting population of various kinds of individuals with a common history or common social, economic or political interests, living together in a common location within a larger society.

Community mobilization The process of organizing and empowering community members to identify problems and resources, facilitate discussion and planning and take action to change their circumstances.

Confidentiality Agreement between client and provider that all conversations will be kept private unless the client gives explicit permission.

Control group A group of persons, facilities or communities that has not been exposed to an intervention. The group, similar to those who received the intervention, provides a comparison and reveals what would have happened had a program not been implemented.

Convenience sampling A method of drawing a sample on the basis of opportunity, for example, grouping youth that are present at a school activity, service providers attending a conference or parents attending a school event.

Counseling The process of providing professional guidance or advice to an individual or group of individuals.

Courtesy bias Distortion of data that occurs when respondents provide less than candid comments because they are influenced by the evaluator (i.e., they know the evaluator or they say what they think the evaluator wants to hear).

Coverage The extent to which those who need something are actually receiving it.

Cross-tabulation A way of presenting, in a table, two or more variables (such as education levels completed by out-of-school youth by gender).

Cultural norm A way of believing or behaving that is typical for a group, based on systems of knowledge, beliefs and relationships in that group.

Data Facts and information collected for a special purpose.

Data collection instruments/methods Tools you use to collect information; the process through which you collect data or information.

Data entry The process of entering data into a computer program prior to analysis.

Demographic survey An instrument that examines and measures a population by a variety of characteristics, such as age, sex, birth and death rates and marital status.

Descriptive statistics Analysis of the general characteristics of a set of data through such measures as frequencies, counts, averages and percentages.

Diagram A visual presentation or drawing that outlines a process or shows the features of an object.

Disaggregate (v.) To separate into parts.

Drop-out One who abandons an attempt, activity or chosen path, such as schooling or participation in a study.

Early marriage Marriage at an age when a person is not physically or psychologically prepared to take on adult roles and responsibilities.

Effectiveness The ability of an intervention to achieve its intended effect in reducing a need or problem in a population.

Efficacy The ability of an intervention to achieve its intended effect in those individuals exposed to it.

Element (as related to sampling) The person or place from which you will collect data; elements can be youth, parents, health facilities or service providers.

Equivalent group Comparison or control groups that are identical to the treatment group except for exposure to a program or intervention, through random assignment.

Ethics The legal, moral and philosophical principles of conduct governing an individual or a group.

Evaluation A process that measures whether program outcomes were achieved and determines what impact the program has had in the target population.

Evaluation design The entire plan of an evaluation, showing methods, timing and how different parts fit together.

Evaluator One who determines the significance, worth or condition of something, usually by careful appraisal and study.

Event log A record of the times, places, participants and other information related to a specific activity.

Executive summary A section, usually at the beginning of a report, that presents a concise overview of the most important and useful findings.

Exit interview A conversation designed to produce feedback from clients after they have participated in a program activity or received a program service.

Experimental design A type of evaluation design that features random assignment to control and experimental groups.

Experimental group A group of individuals assigned to receive a given treatment or intervention.

Extraneous events Events or conditions external to the intervention that might influence project participants and their outcomes.

Facility A place, such as a hospital or clinic, that is built, installed or established to serve a particular purpose.

Family life education (FLE) An intervention that teaches young people in and out of school about adult (marital and parental) roles and expectations, marriage and childbearing, relationships, anatomy, physiology and, in some cases, pregnancy and STI/HIV prevention.

Feedback Subjective information given to the project about an action, event or process.

Fertility The birthrate of a population or the capacity of individuals to reproduce.

Field notes Formal and informal records of a data collector's observations, thoughts and conversations during data collection.

Flow chart A graphic design that shows the separate parts of a larger whole and the ways they are linked to one another.

Focus group discussion A data collection method that identifies issues, terms and interpretations from a group of individuals with similar characteristics.

Follow-up survey A survey initiated after intervention activities have begun to measure changes during a designated time period.

Free listing A technique used to uncover the ways people understand and interpret a particular subject by analyzing local terms, their relative importance and their relationships to one another.

Frequency A univariate (single variable) measure used to summarize a number of observations, for example, the education levels completed by out-of-school youth.

Full-coverage program A program or intervention that is expected to reach the entire target population of young adults (e.g., intensive mass media compaigns and national family life education programs).

Gender Social and cultural definitions of masculine and feminine roles, identities and character traits.

Generic control The use of data on the general population from a non-program source as a comparison population.

Goal A description of the overall impact expected of a program.

Graph A presentation of data featuring characteristics on a horizontal and vertical axis.

Health The general condition of being of sound body, mind or spirit, especially the freedom from physical disease or pain.

Health education The process of providing information that builds skills in ways that promote individual and community health.

Health facility A place where health services are provided, such as a hospital, clinic or pharmacy.

Health-seeking behavior Actions or responses by individuals to protect their health.

Impact An effect or change directly caused by something external, such as a change in an individual caused by his or her participation in a program.

Impact analysis Identification of patterns in data that show changes in desired outcomes that can be attributed to an intervention.

Impact evaluation A type of evaluation that determines how much of an observed change in outcomes is due to a program's efforts.

Implementation The process of carrying out program activities.

Incidence The rate of occurrence of something in a population.

Inconsistent data Responses to one question that are not consistent with responses to other, similar questions.

Indicator A measurable statement of program objectives and activities. It may be expressed in numeric or non-numeric terms, and express quantitative or qualitative factors.

Inferential statistics Analysis of data that allows the evaluator to draw conclusions about the population from which the sample data were drawn, based on probabilities. Inferential statistics assess the likelihood of an event occurring.

Informal listening session A less structured and more spontaneous group discussion that produces information about expected cultural norms.

Informant Someone who is interviewed by an evaluator.

Institution An established organization or corporation.

Instrument A tool (e.g., form, questionnaire or checklist) used to collect information or data.

Intermediate outcome The midterm result a program hopes to produce, usually related to changes in behavior.

Intervention An activity that aims to maintain or alter the condition of those it reaches.

Interview An intensive, one-to-one exchange.

Inventory A data collection method that assesses the services provided with regard to quality and quantity of facilities, equipment and supplies.

Key element A factor or component of a program that is regarded as important to the program's success.

Key informant A respondent who has special knowledge, status or access to observations that are important to the program and who is willing to share his or her knowledge and skills.

Level of confidence The degree of certainty that an observed change is due to the intervention rather than to random chance.

Level of significance The probability with which you want to be certain that the magnitude of change did not occurr by chance.

Line graph A drawing that connects points on a graph, using straight lines.

Linked sampling A method whereby all sample elements (e.g., students, parents, teachers and community leaders) are included in the same sample.

Logic Model An organized planning sequence of defining goals, identifying antecedent factors and selecting program activities to influence the antecedents.

Longitudinal study A study that follows the same group of participants over an extended time period.

Long-term outcome An outcome that a program hopes to achieve after a longer period of operation, usually related to observations at the population level.

Macro level A level of analysis that refers to the interrelationship of large sectors, or things that occur on a large scale.

Magnitude of change The difference between baseline data and data collected at the second round of data collection.

Management information system (MIS) A framework set up to systematically compile and maintain program information.

Mapping A structured activity where individuals or groups make a graphic representation of either a part or the whole of their residential, work or social environment.

Mass media Communication that reaches large numbers of people in a population, such as television, movies, advertising, radio and newspapers.

Matching The process of selecting subjects (in the treatment and control groups) that share common characteristics, such as age, gender, education level or location.

Mean The mathematical average of series of numbers (the sum of all scores divided by the number of cases).

Measure To examine the extent or quantity of something by comparing it with a fixed unit or object of known size.

Measure of size (MOS) A count or estimate of the number of sample elements associated with each cluster.

Method A way, technique or process of doing something.

Methodology A particular procedure or set of procedures.

Missing data The result of a respondent declining to answer a question, a data collector failing to ask or record a respondent's answer or a data entry staff member skipping the entry of a response.

Monitor The systematic and continuous following, or keeping track, of activities to ensure that they are proceeding according to plan.

Morbidity The relative incidence of a disease.

Mortality The number of deaths in a given time or place.

Multi-stage sampling A type of sampling method with several stages of sampling and sub-sampling, usually used in large-scale population surveys.

Mystery client A trained person who visits program facilities in the guise of a client, systematically recording his or her experiences after the service encounter.

Network A pattern of formal or informal links between individuals, organizations and other sources of information, or a group or system of interconnected or cooperating individuals.

Non-equivalent groups Treatment and control groups that are similar but not necessarily identical, with regard to all factors other than exposure to a program or intervention.

Non-experimental design A study design that does not use control groups for comparison with the intervention groups.

Non-probability method The process of choosing samples by convenience sampling, quota sampling and snowball sampling rather than random chance.

Non-response Instances in which a sample subject was not interviewed due to refusal or inability to contact that person.

Nutrition The act or process of being nourished through the consumption of food.

Objective A measurable statement of program outcomes.

Observation techniques Systematic evaluation methods for observing people, events and/or their contexts.

Open-ended questions Queries that allow respondents to answer in whatever way they wish, without determined response categories.

Outcome evaluation A type of evaluation that determines whether selected outcomes have changed in the target population.

Out-of-range values Data items that seem impossible or implausible that may be a result of a data entry error or the respondent misunderstanding the question.

Outreach A strategy that reaches young people in the places where they congregate, as opposed to requiring them to visit a site for information or services.

Panel study A study that follows a group of persons or communities and collects data from them at several points in time.

Participant An individual who takes part in a service, program or activity.

Participatory evaluation A systematic evaluation process that promotes stakeholders' participation in initial decision making, planning, implementation, summary, analysis and use of results.

Partner A person who is romantically and/or sexually associated with another.

Peer A person that is of equal standing to another, often of the same age, economic background and educational level.

Peer education/peer promotion Programs that feature young people educating or promoting services among their peers.

Percentage The number of people with a particular characteristic in a group, divided by the total number in the group and multiplied by 100.

Performance target The intended result to be achieved within a stated time.

Pie chart A circular chart used to show the different parts of a whole in relation to one another.

Pile sorting A data collection method used to cluster items into different categories.

Policy A plan that spells out the general goals and acceptable procedures of a government or organization.

Population census A complete count or enumeration of the population of a country or other geographical region, usually undertaken every 10 years.

Population-level objective A measurable statement—in terms of the target population—of the impact a program hopes to have.

Population survey Highly structured surveys, carried out with some form of probability sampling in the program's target population.

Pornography The sexually explicit subordination of a person, graphically depicted in pictures or with words.

Posttest A test after an intervention that measures the participants' knowledge and skills, which can be compared to what they knew before the intervention.

Power The probability with which you want to be certain of detecting the magnitude of change if one actually occurred.

Precision The extent to which a survey estimate differs from the "true" value of an indicator, due to random error.

Pre-coded responses Items included in data collection instruments for which all possible responses are listed.

Press release A concise statement to the news media that presents news, such as an overview of evaluation findings or other information.

Pretest A test before an intervention that measures the participants' knowledge or skills, which can be compared to what they know after the intervention.

Primary data New information that has been collected through surveys, qualitative community research, etc.

Privacy Freedom from unwanted intrusion.

Probability The likelihood that an event will occur.

Probability-proportional-to-size (PPS) In sampling, a method that gives larger clusters a greater chance of selection than smaller clusters to compensate for the fact that elements in large clusters have a lower chance of being chosen at the second stage of sample selection than elements in smaller clusters.

Probability sampling methods Types of sampling methods that are based on probability theory, a mathematical concept that refers to the ability to predict the statistical likelihood that a random event will occur.

These sampling methods include simple random, systematic, stratified, cluster and multi-stage.

Process evaluation Evaluations that are carried out while a program is underway to measure how well program activities are performed.

Program design The process of developing a program in response to a defined community problem or need.

Program-level objective A measurable statement of the structure, management or operations of a program, describing the steps that need to be taken to achieve the intended impact.

Program outcome The specific result that a program hopes to achieve.

Project (v.) To forecast how an indicator will change over time.

Proportion A frequency divided by the total number of cases. The numerator is a portion of the total, the denominator is the total number of cases.

Provider An individual who provides a service, such as a doctor, paramedic, nurse or counselor.

Qualitative Non-numeric data or indicators that are expressed in words.

Qualitative methodology A set of procedures used to collect qualitative data, such as focus group discussions, body mapping and free-listing.

Quality The degree of excellence or standard of an entity.

Quantitative Measured by or concerned with amount or quantity, and expressed in numbers or quantities.

Quasi-experimental design A study design in which similar experimental groups are assigned through non-random methods. A control group that is similar to the experimental group is often chosen by matching characteristics.

Questionnaire A group of written or printed questions used to obtain information from individuals or groups.

Quota sampling A type of sampling method where groups are defined in advance of data collection and a sample is chosen for each group, but not necessarily at random.

Random Having the same probability of occurring as every other member of a set.

Random assignment The process of assigning persons, facilities or communities to either the treatment or control group at random (for example, by flipping a coin or drawing numbers).

Random sample Selection of members of a population (or other items) in such a way that everyone (or everything) has an equal chance of being included.

Randomized experiment A type of study design where subjects are assigned by chance (i.e., randomly) to treatment and control groups prior to program implementation.

Ranking A data collection method used to organize information by preference or relative importance.

Reliability The extent to which a survey estimate differs from the "true" value of an indicator due to random error.

Religiosity The degree to which an individual's religious faith shapes his or her beliefs, values and behavior.

Representative (adj.) The extent to which data collected accurately describe the target population for a program.

Reproductive health The health and well-being of women and men in terms of pregnancy, birth and related conditions, diseases and illnesses.

Reproductive health outcome Sexual health changes observed at the population level among members of the target population as a result of a given program or intervention, such as delay of sexual initiation, prevention of unintended pregnancies or promotion of breastfeeding.

Research The careful, systematic study or investigation in a particular subject area.

Resource A source of information or expertise.

Respondent A person who provides information during a data collection effort, for example, by completing a questionnaire.

Result A consequence or conclusion of a project.

Retrospective evaluation A study in which the previous histories of people are examined for exposure to a potentially causative agent.

Role play (n.) A dramatic enactment of a real-life situation.

Sample A part of a whole selected to represent that whole (for example, a sample of the population).

Sample size The number of sample elements from which you will need to collect data in order for your evaluation findings to be statistically significant.

Sample survey A survey of a sample of a specific population.

Sampling The process and techniques of studying part of something to gain information about the whole (like a population).

Sampling element The unit from or about whom data is to be collected. A sample element could be a young person, a parent or a program site, such as school or a health center.

Sampling frame A list of all relevant sample elements that is used to select a sample. Sampling frames should include all elements in the target population.

Scope The extent of activity undertaken in a monitoring and evaluation effort.

Secondary data Data that are already available through recent surveys, qualitative community research or administrative reports.

Selection bias Distortion of data that arises from inherent differences between experimental and comparison groups that might account for their posttest differences on outcomes.

Sensitivity The extent to which a method gives results that are free from false negative measurements.

Service delivery The different components or operations offered to clients, such as training, clinical services, counseling, education or commodities.

Service statistics Program information usually compiled in the form of counts that provide a quantitative description of program activities undertaken, such as numbers of events or number of clients.

Sexual health The health and well-being of women and men in terms of sexuality and related conditions, diseases and illnesses.

Sexuality education The provision of information about sexual and reproductive health, and the strengthening of decision-making skills related to healthy sexual and reproductive health.

Short-term outcome The specific outcome a program hopes to produce in a small period of time, usually related to changes in the risk and protective factors that are related to decision making about behavior.

Significance The determination that differences observed in data, however small or large, were not due to chance.

Simple random sampling A type of sampling method where elements are chosen at random so that each element has an equal chance of selection.

Site The location of a program (e.g., school, clinic or community).

Site visit A prearranged visit by supervisors or evaluators to program locations, usually in order to conduct monitoring or process evaluation.

Snowball sampling A method of sampling a target group by asking a small group of people with special characteristics to identify other people like them. This process continues until a target sample size has been reached, or until additional data collection yields no new information.

Social context The background, environment or situation relating to relationships among individuals, groups and institutions.

Social mapping A data collection method that reveals social and physical boundaries, the social infrastructure, housing patterns or other features that are relevant to social interactions.

Social marketing A process of promoting or selling ideas, products or values that contribute to improvements in the health or social welfare of a group or population.

Social system The set of relationships, rules and regulations in a society that governs behavior.

Socio-demographic Pertaining to characteristics of an individual's, group's or population's gender, age, education, income or economic status, place of residence, marital status, ethnicity and employment.

Socio-drama An enactment of a situation that occurs in a culture or society to educate or change values and behaviors; similar to a role play.

Staff The people who work for a program.

Stakeholders Persons outside the immediate program staff who have an interest and role in program functions and activities.

Statistical associations A statistical relationship between two indicators or variables.

Statistical significance The determination that experimental results are not likely to be due to chance, or are not attributable to fluctuations associated with random sampling procedures.

Statistics Numerical facts that are systematically collected, organized and presented in a special way.

Strategy A plan developed to achieve a specified goal, objective or result.

Stratification The process of dividing the target population into sub-groups and then taking a sample from each.

Stratified sampling study design A type of sampling method where the population to be sampled is divided into homogenous groups, based on characteristics you consider important to the indicators being measured. A simple random or systematic sample is then chosen from each group.

Survey An instrument that collects factual information from multiple respondents.

Survey methods The techniques and procedures used to obtain information, often about a human population.

Systematic sampling A type of sampling method where the first element is chosen at random. Subsequent elements are chosen using a fixed interval (e.g., every tenth element) until you reach the desired sample size.

Systems development and functioning Activities programs need to support implementation, which are carried out before and during the work with program participants.

Table A graphic presentation of facts and numbers in an orderly fashion, usually in columns and rows, so that they can be more easily understood.

Tabulate To count, record or list systematically.

Tally sheet A list that enumerates key features of a setting or process that users can check off during an observation; a checklist.

Target The level of an objective that one plans to achieve within a stated time.

Target population The specific group of individals a program is trying to affect, influence or serve.

Theory A statement of apparent relationships and underlying principles used to explain behaviors, acts or events.

Time-series design An approach of taking measurements of indicators both before and after the intervention is implemented to detect program impact as changes in trends after the intervention was introduced.

Tool An instrument used to collect information, such as a form or survey.

Transcribe To write out verbatim a tape-recorded conversation.

Transcript Written verbatim records of what was said, for example, between an interviewer and respondents during a focus group discussion or interview.

Treatment group Persons, facilities or communities that are exposed to a program or intervention.

Trend data Data that are collected at multiple points in time in order to monitor changes.

Utilization The practical use of materials, services or information.

Validity The ability of a study design to measure the "true" impact of a program or intervention.

Values An individual's or group's beliefs and opinions that, together with personal characteristics, group norms, etc., often determine behavior.

Variables The assorted characteristics of an entity that is being looked at or measured.

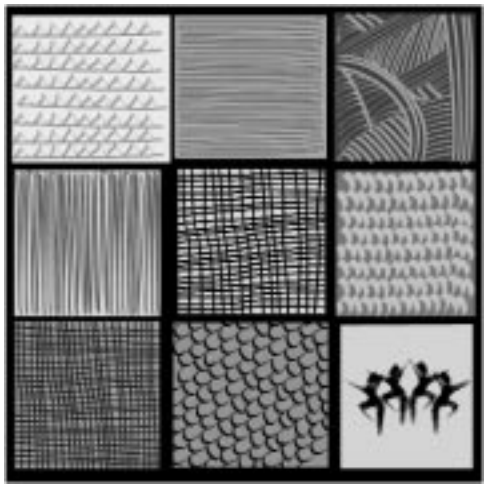
Vital Registration System A data-gathering procedure for compiling information on births and deaths.

Vocational training Education to develop the skills needed to pursue a particular occupation or type of work.

Workplan A detailed outline of what activities will take place in order to achieve specified outcomes and objectives.

Youth A young person in the transition between childhood and adulthood, commonly defined as between the ages of 10 and 24.

PHOTO: JHU/CCP



BIBLIOGRAPHY

Amazigo, U., N. Silva, J. Kaufman and D.S. Obikeze. 1997. Sexual activity and contraceptive knowledge and use among in-school adolescents in Nigeria. *International Family Planning Perspectives* 23 (1): 28–33.

Amon, J., K. Kotellos and W. Githens Benazerga. December 1997. *Capacity Building.* Arlington, VA: Family Health International, The AIDS Control and Prevention (AIDSCAP) project.

BASICS. Spring 1998. "Child survival BASICS." *Basic Support for Institutionalizing Child Survival, Project Number 5.* Arlington, VA: BASICS.

Bassett, M. and J. Sherman. 1994. "Female sexual behavior and the risk of HIV infection: An ethnographic study in Harare, Zimbabwe." *Women and AIDS Program Research, Report Series No. 3.* Washington, DC: International Center for Research on Women (ICRW).

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

Bertrand, J.T. and D.L. Kinkaid. October 1996. *Evaluating Information-Education-Communication (IEC) Programs for Family Planning and Reproductive Health, Final Report of the IEC Working Group.* Chapel Hill, NC: The EVALUATION Project.

Bertrand, J.T., R.J. Magnani and J.C. Knowles. 1994. *Handbook of Indicators for Family Planning Program Evaluation.* Chapel Hill, NC: The EVALUATION Project.

Bertrand, J.T., R.J. Magnani, and N. Rutenberg. September 1996. *Evaluating Family Planning Programs with Adaptations for Reproductive Health.* Chapel Hill, NC: The EVALUATION Project.

Bertrand, J.T., R. McCloskey, M.A. Hoskins and E. Mielke. April 1994. *Inventory of Indicators Used to Evaluate Adolescent Reproductive Health Programs in Latin American and the Caribbean (Draft).* Washington, DC: Center for Population Options.

Bertrand, J.T., J. Stover and R. Porter. n.d. "Methodologies for evaluating the impact of contraceptive social marketing." *Practical Guide No. 5.* Washington, DC: SOMARC/The Futures Group International.

Birdthistle, I. and C. Vince-Whitman. 1997. *Reproductive Health Programs for Young Adults: School-Based Programs.* Washington, DC: FOCUS on Young Adults.

Blair, J. August 1995. "Impact On Teen Risk Reduction." (Unpublished report prepared for Population Services International, PSI/Project ACTION, Washington, DC).

Blum, R. 1998. *Reducing the Risk: Connections that Make a Difference in the Lives of Youth.* Bethesda, MD: Add Health.

Blum, R. 1999. "Youth Health and Development: Conceptual Issues and Measurement." (Presentation to the WHO Meeting on Adolescent Health and Development, Washington, DC, 4–6 February 1999).

Blumenfeld, S.N. May 1985. *Operations Research Methods: A General Approach in Primary Health Care.* PRICOR Monograph Series: Methods Paper 1. Washington, DC: Center for Human Services.

Bongaarts, J. and S.C. Watkins, December 1996. Social interactions and contemporary fertility transitions. *Population Development Review* 22 (4).

Boohene, E., J. Tsodzai, K. Hardee-Cleaveland, S. Weir and B. Janowitz. 1991. Fertility and contraceptive use among young adults in Harare, Zimbabwe. *Studies in Family Planning* 22 (4): 264–271.

Brindis, C. and L. Davis. 1998a. *Communities Responding to the Challenge of Adolescent Pregnancy Prevention: Volume 2, Building Strong Foundations, Ensuring the Future.* Washington, DC: Advocates for Youth.

Brindis, C. and L. Davis. 1998b. *Communities Responding to the Challenge of Adolescent Pregnancy Prevention. Volume 4: Improving Contraceptive Access for Teens.* Washington, DC: Advocates for Youth.

Brindis, C., J.L. Peterson, J.J. Card and M. Eisen. 1996. *Prevention Minimum Evaluation Data Set: A Minimum Data Set For Evaluating Programs Aimed at Preventing Adolescent Pregnancy and STD/HIV/AIDS, 2nd Edition.* Los Altos, CA: Sociometrics Corporation.

Brindis, C., K. Pittman, R. Reyes and S. Adams-Taylor. 1991. *Adolescent Pregnancy Prevention: A Guidebook for Communities.* Palo Alto, CA: Stanford Health Promotion Resource Center.

Brinkerhoff, R. O., D. M. Brethower, T. Hluchyj, J. R. Nowakowoski. 1983. *Program Evaluation: A Practitioner's Guide for Trainers and Educators.* Boston, MA: Kluwer-Nijhoff Publishing.

Bruce, J. 1990. Fundamental elements of quality care: A simple framework. *Studies in Family Planning* 21 (2): 61–90.

Buckner, B., A.O. Tsui, A.I. Hermalin and C. McKaig, eds. January 1995. *A Guide to Methods of Family Planning Program Evaluation, 1965–1990.* Chapel Hill, NC: The EVALUATION Project.

Burt, M.R., A.V. Harrell, L.C. Newmark, L.Y. Aron and L.K. Jacobs. December 1997. *Evaluation Guidebook for Projects Funded by S.T.O.P. Formula Grants under the Violence Against Women Act.* Washington, DC: The Urban Institute.

Campbell, D.T. and J.C. Stanley. 1966. *Experimental and Quasi-Experimental Designs for Research.* Chicago, IL: Rand McNally.

Caldwell, J.C., B.K. Caldwell and I. Pieris. 1998. "The Construction of Adolescence in a Changing World: Implications for Sexuality, Reproduction and Marriage." *Studies in Family Planning* 29 (2): 137–153.

Card, J.J., S. Niego, A. Mallai and W.S. Farrell. 1996. The program archive on sexuality, health and adolescence: Promising "prevention programs in a box." *Family Planning Perspectives* 28 (5): 210–220.

Chege, I., J. Aluar and A. Ngay. 1995. "Executive summary." *Final Evaluation Report of the Communication Resources for the Under 18s on STDs and HIV (Crush) Project.* Nairobi: CARE Kenya.

Choi, K. and T. Coates. 1994. Prevention of HIV infection. *AIDS* 8: 1371–89.

David, H.P. and J. Rademakers. 1996. Why Dutch teenagers have the low rates for pregnancies and abortion. *Studies in Family Planning* 27 (6): 341–43.

Debus, M. 1988. *Methodological review: A handbook for excellence in focus group research.* (Unpublished guide prepared for Academy for Educational Development HEALTHCOM Project, Washington, DC).

Drolet, J.C. and K. Clark, eds. 1994. *The Sexuality Education Challenge: Promoting Healthy Sexuality in Young People.* Santa Cruz, CA: ETR Associates.

Eisen, M. and G.L. Zellman. 1992. "A health beliefs field experiment." In Miller, B.C., J.J. Card and R.L. Paikoff, eds. *Preventing Adolescent Pregnancy: Model Programs and Evaluations.* Newbury Park, CA: Sage Publications.

Family Health International and FOCUS on Young Adults Program. 1997. *Reproductive health of young adults: Contraception, pregnancy and sexually transmitted diseases. Contraceptive Technology Update Series.* Chapel Hill, NC: Family Health International.

Fawcett, S.B. 1995. *Work Group Evaluation Handbook: Evaluating and Supporting Community Initiatives for Health and Development.* Lawrence, KS: Work Group on Health Promotion and Community Development.

Fawcett, S.B., et al. In press. "Evaluating community initiatives for health and development." In Rootman, I., et al., eds. *Evaluating Health Promotion Approaches.* Copenhagen, Denmark: World Health Organization.

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

Fetro, J.V. 1994. "Evaluating sexuality education programs." In Drolet, J.C. and K. Clark, eds. *The Sexuality Education Challenge: Promoting Healthy Sexuality in Young People.* Santa Cruz, CA: ETR Associates.

Feuerstein, M.T. 1994. *Partners in Evaluation: Evaluating Development and Community Programmes with Participants.* London and Basingstoke: The MacMillan Press Ltd.

Fisher, A., J. Laing, J. Stoeckel and J. Townsend. 1991. *Handbook for Family Planning Operations Research Design, 2nd Edition.* New York: The Population Council.

Fisher, A., B. Mensch, R. Miller, I. Askew, A. Jain, C. Ndeti, L. Ndhlovu and P. Tapsoba. 1992. *Guidelines and Instruments for a Family Planning Situation Analysis Study.* New York, NY: The Population Council.

FOCUS on Young Adults Program. July 1998. *Monthly Update.* Washington, DC: FOCUS on Young Adults.

FOCUS on Young Adults and the Program of Family and Sexual Education, Ministry of Education, Peru. 1999. *Evaluation Results 1996–1998 and Baseline Survey for the 1999–2000 Program Period.* Washington, DC and Lima, Peru: FOCUS on Young Adults and Pathfinder, International, Peru.

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

Frost, J.J. and J.D. Forrest. 1995. Understanding the impact of effective teenage pregnancy prevention programs. *Family Planning Perspectives* 27 (5): 188–95.

Gage, A. June 1998. Sexual activity and contraceptive use: The components of the decision-making process. *Studies in Family Planning* 29 (2).

García-Núñez, J. 1992. *Improving Family Planning Evaluation.* West Hartford, CT: Kumarian Press.

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

Grosskurth, H., F. Mosha, J. Todd et al. 1995. Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: Randomised controlled trial. *Lancet* 346: 530–36.

Hanson, G.A., K.T. Hanson and T.D. Stoddard. 1995. *Say It Right: A Guide to Effective Business Presentations.* Chicago, IL: Irwin Publishers.

Hardee, K. and K.M. Yount. 1995. *From Rhetoric to Reality: Delivering Reproductive Health Promises Through Integrated Services.* Research Triangle, NC: Family Health International, Women's Study Project.

Health Technical Services. Spring 1998. *HTS in Brief,* Number 4.

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

Houvras, I. and C. Kendall. 1997. *Influencing Young Adult Reproductive/Sexual Behaviors: A Review of the Literature.* (Unpublished paper prepared for the FOCUS on Young Adults, Tulane University, New Orleans, LA).

Hughes, M.E., F.F. Furstenberg and J.O. Teitler. 1995. The impact of an increase in family planning services on the teenage population of Philadelphia. *Family Planning Perspectives* 27 (2): 60–65.

Infante, J.M., M. Diaz, J. Benavides and C. Brambila. n.d. *Institutionalization of Youth Programs: A Follow-up Survey.* Monterrey, NL, Mexico: Pro-Superacion Familiar Neoleonesa.

Israel, R. and M. Nagano. 1997. *Reproductive Health Programs for Young Adults: Social Marketing and Mass Media Approaches.* Washington, DC: FOCUS on Young Adults.

Jimerson, A.B. and D. Stone. n.d. "HIV prevention in the schools." In Academy for Educational Development. *A World Against AIDS: Communication for Behavior Change.* Washington, DC: Academy for Educational Development/AIDSCOM.

Kauffman, C., J. Brower, L. Hue and S. Randolph. 1996. *Island-Wide HIV/STD Prevention Program.* (AIDSCAP Subproject Final Narrative Report, Jamaica Red Cross, American Red Cross, October 21, 1996).

Kirby, D. 1997. *No Easy Answers: Research Findings on Programs to Reduce Teen Pregnancy.* Washington, DC: The National Campaign to Prevent Teen Pregnancy, Task Force on Effective Programs and Research.

Kirby, D. 1999a. "Antecedents of Sexual Behaviors and Pregnancy." (Presentation to the WHO meeting on Adolescent Health and Development, Washington, DC, 4–6 February 1999.)

Kirby, D. 1999b. *The Multiple Antecedents of Adolescent Sexual Risk-Taking Behavior, Pregnancy and Childbearing, and their Implicaitons for Research and Programs.* (Draft paper prepared with support from the California Wellness Foundation).

Kirby, D. 1999c. Reducing adolescent pregnancy: Approaches that work. *Contemporary Pediatrics* 16(1): 83–93.

Kirby, D.B. and L. Brown. 1996. Condom availability programs in U.S. schools. *Family Planning Perspectives* 28 (5): 196–202.

Kirby, D., M. Korpi, R. Barth and H. Cagampang. 1997. The impact of the *Postponing Sexual Involvement* curriculum among youth in California. *Family Planning Perspectives* 29: 100–108.

Kirby, D., L. Short, J. Collins, D. Rugg, L. Kolbe, M. Howard, B. Miller, F. Sonenstein and L.S. Zabin. 1994. School-based programs to reduce sexual risk behaviors: A review of effectiveness. *Public Health Reports* 109 (3): 339–360.

Kirkpatrick, J.S. 1981. *The Magic of Structure: A Manual for Planning and Evaluating Family Planning Education Programs.* New York, NY: Planned Parenthood Federation of America, Department of Education.

Kolbe, L. 1990. An epidemiological surveillance system to monitor the prevalence of youth behaviors that most affect health. *Health Education* 21: 44–47.

Koo, H.P., G.H. Dunteman, C. George, Y. Green and M. Vincent. 1994. Reducing adolescent pregnancy through a school- and community-based intervention: Denmark, South Carolina, revisited. *Family Planning Perspectives* 26 (5): 206–11.

Kotellos, K., K. Hartwig and W. Githens. 1994. *A Framework for Incorporating Evaluation into Project Design.* Module 3, AIDSCAP Evaluation Tools. Washington, DC: AIDS Control and Prevention Project (AIDSCAP), Family Health International.

Krause, D. 1996. *Effective Program Evaluation, an Introduction.* Chicago, IL: Nelson-Hall Publishers.

Labovitz, S. and R. Hagedorn. 1971. *Introduction to Social Research.* New York, NY: McGraw-Hill Book Company.

Lawrence, J.E.S. 1989. Engaging recipients in development evaluation: The "stakeholder approach." *Evaluation Review* 13 (3): 243–56.

Leffert N, P. Benson, P. Scales, A. Sharma, D. Drake and D. Blyth. 1998. Developmental assets: Measurement and prediction of risk behaviors among adolescents. *Applied Developmental Science* 2 (4): 209–30.

MacLaren, L. 1997. *Monitoring and Evaluation: Lessons Learned from the Lentera Project.* London: International Planned Parenthood Federation.

Maher, S., S. Smith and C. Sorensen. June 1995. *Working Group on the Evaluation of Family Planning Management: Final Report.* Chapel Hill, NC: The EVALUATION Project, Carolina Population Center, University of North Carolina at Chapel Hill.

McCauley, A.P. and C. Salter. 1995. Meeting the needs of young adults. *Population Reports* Series J (41). Baltimore, MD: Johns Hopkins School of Public Health, Population Information Program.

McKaig, C., et al. 1996. *Adolescent Reproductive Health Programs in Developing Countries: A Review of the Literature.* Washington, DC: FOCUS on Young Adults.

Meekers, D., G. Stallworthy and J. Harris. 1997. Changing adolescents' beliefs about protective sexual behavior: The Botswana Tsa Banana program. *Working Paper No. 3.* Washington, DC: Population Services International.

Mensch, B., J. Bruce and M. Greene. 1998. *The Unchartered Passage: Girls' Adolescence in the Developing World.* New York: The Population Council.

Merritt, A.P. and M. Raffaelli. 1993. Creating a model HIV prevention program for youth. *The Child, Youth, and Family Services Quarterly* 16 (2).

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

Miller, B.C., J.J. Card, R.L. Paikoff, J.L. Peterson, eds. 1992. *Preventing Adolescent Pregnancy: Model Programs and Evaluation.* Newbury Park, CA: Sage Publications.

Miller, D.C. 1977. *Handbook of Research Design and Social Measurement.* New York: David McKay Company, Inc.

Miller, J. and C. Bahamon. 1995. *Family Planning Management Terms. A Pocket Glossary in Three Languages.* Boston, MA: Family Planning Management Development, Management Sciences for Health.

Miller, R., A. Fisher, K. Miller, L. Ndhlovu, B.N. Maggwa, I. Askew, D. Sanogo and P. Tapsoba. 1997. *The Situation Analysis Approach to Assessing Family Planning and Reproductive Health Services.* New York: The Population Council, Africa Operation Research and Technical Assistance Project II.

Millstein, S., A. Petersen and E. Nightingale. 1993. *Promoting the Health of Adolescents.* New York, NY: Oxford University Press.

Monroy, A., C. Esteve, R. Lundgren, S. Rosenhouse, and J. Townsend, 1987. "Prospective Cost-Effectiveness Study to Determine a Strategy of Expansion of Services to Young Adults in Mexico City." (Final Technical Report, Centro de Orientacion para Adolescentes (CORA). October 1, 1986–July 31, 1987).

Monroy, A., C. Esteve, R. Lundgren, S. Rosenhouse and J.W. Townsend. 1988. "Prospective Cost-Effectiveness Study to Determine a Strategy of Expansion of Services to Young Adults in Mexico City." (Final Technical Report. Mexico, DF: Centro de Orientacion para Adolescentes (CORA) and the Population Council).

Moore, K.A., B.C. Miller, D. Glei and D.R. Morrison. 1995. *Adolescent Sex, Contraception, And Childbearing: A Review of Recent Research.* Washington, DC: Child Trends, Inc..

Moore, K.A., B.W. Sugland, C. Blumenthal, D. Glei and N. Snyder. 1995. *Adolescent Pregnancy Prevention Programs: Interventions and Evaluations.* Washington, DC: Child Trends, Inc.

Mullen, P.D. and D.C. Iverson. 1986. "Qualitative methods." In Green, L. and F.M. Lewis, eds. *Measurement and Evaluation in Health Education and Health Promotion.* Palo Alto, CA: Mayfield Publishing Company.

Nelson, K., L. MacLaren and R. Magnani. 1999. *Assessing and Planning for Youth Friendly Reproductive Health Services.* Washington DC: FOCUS on Young Adults.

Pathfinder International. 1998. *Insights from Adolescent Project Experience, 1992–1997.* Watertown, MA: Pathfinder International.

Pathfinder International, Evaluation Unit. 1995. *Adolescent Project Evaluation (Draft).* Watertown, MA: Pathfinder International.

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

Paxman, J.M. and R.J. Zuckerman. 1987. *Laws and Policies Affecting Adolescent Health.* Geneva: World Health Organization.

Peace Corps. 1981. *Assessing Performance: A Reference Series for the Field. Project Evaluation Handbook, Volume 1.* Washington, DC: Peace Corps Evaluation, Office of Personnel and Management.

Perry, C.L. and R. Sieving. 1991. "Involvement in Global AIDS Prevention Among Adolescents." (Unpublished review commissioned by the Global Programme on AIDS, World Health Organization, Geneva).

Philliber, S. and J.P. Allen. 1992. "Life options and community service." In Miller, B.C., J.J. Card, R.L. Paikoff and J.L. Peterson, eds. *Preventing Adolescent Pregnancy: Model Programs and Evaluations.* Newbury Park, CA: Sage Publications.

Pick de Weiss, S. and P. Andrade Palos. 1989. *Development and Longitudinal Evaluation of Comparative Sex Education Courses.* Mexico, DF: Instituto Mexicano de Investigacion de Familia y Poblacion.

Pick de Weiss, S., P. Andrade Palos and J. Townsend. 1990. *Planeando Tu Vida: Development and Testing of a Family Life Education Program for Young Adults, Final Technical Report.* Mexico, DF: Instituto Mexicano de Investigacion de Familia y Poblacion.

Planned Parenthood Association of Ghana. 1997. *Adolescent Reproductive Health Information and Services, Baseline Survey: Interviewers' Training Manual.* Ghana: Vision 2000 Project, Volta Region.

The POLICY Project. 1998. "Program Design and Evaluation Guidelines." (Unpublished document prepared by The Futures Group International.)

Popham, W.J. n.d. *Evaluating HIV Education Programs.* Atlanta, GA: Division of Adolescent and School Health, Centers for Disease Control and Prevention.

Rappaport, J. 1984. Studies in empowerment: Introduction to the issue. *Prevention in Human Services* 3 (2,3): 1–7.

Resnick, M, et al. September 1997. Protecting adolescents from harm: Findings from the national longitudinal study on adolescent health. *The Journal of the American Medical Association* 278 (10).

Rimon, J.G., K.A. Treiman, D.L. Kincaid, A. Silayan-Go, M.A. Camacho-Reyes, R.M. Abejuela, and P.L. Coleman. November 1994. *Promoting Sexual Responsibility in the Philippines Through Music: An Enter-Educate Approach.* Occasional Paper Series, No. 3. Baltimore, MD: Johns Hopkins School of Public Health, Center for Communication Programs.

Rossi, P. and H. Freeman. 1993. *Evaluation: A Systematic Approach, 5th Edition.* Newbury Park, CA: Sage Publications.

Samuels, S.E. and M.D. Smith. 1993. *Condoms in the Schools.* Menlo Park, CA: Henry J. Kaiser Family Foundation.

Scrimshaw, S.C.M. and E. Hurtado. 1987. *Rapid Assessment Procedures for Nutrition and Primary Health Care.* Tokyo: The United Nations University, and Los Angeles, CA: University of California.

Seligman, B., N. Jewell, and J. Paxman, 1996. "Creating a Favorable Policy Environment for Young Adult Reproductive Health: A Legal and Regulatory Framework." (Paper presented to the American Public Health Association meeting, New York).

Senderowitz, J. September 1996. *Thematic Evaluation on Adolescent Reproductive Health, Global Report (Draft).* New York: UNFPA.

Senderowitz, J. 1997a. *Reproductive Health Programs for Young Adults: Health Facility Programs.* Washington, DC: FOCUS on Young Adults.

Senderowitz, J. 1997b. *Reproductive Health Programs for Young Adults: Outreach Programs.* Washington, DC: FOCUS on Young Adults.

Senderowitz, J. 1998. *Involving Youth in Reproductive Health Programs.* Washington, DC: FOCUS on Young Adults.

Senderowitz, J. 1999. *Making Reproductive Health Services Youth Friendly.* Washington, DC: FOCUS on Young Adults.

Shah, M.K., R. Zambezi and M. Simasiku. 1999. *Listening to Young Voices: Facilitating Participatory Appraisals with Adolescents on Reproductive Health.* Washington, DC: Care International in Zambia and FOCUS on Young Adults.

Shah, M.K., R. Zambezi and M. Simasiku. 1998. *The Use of PLA Methodology in Partnership for Adolescent Sexual and Reproductive Health Project (PALS): A Field Guide.* Washington, DC: Care International in Zambia and FOCUS on Young Adults.

Solomon, B. 1993. "Legal Issues." In Samuels, S.E. and M.D. Smith, eds. *Condoms in the Schools.* Menlo Park, CA: Henry J. Kaiser Family Foundation.

Speizer, I. and S. Mullen. 1997. *Literature Review on Adolescent Reproductive Health.* (Unpublished paper prepared for FOCUS on Young Adults, Washington DC).

Stewart, L. and E. Eckert, eds. December 1995. *Indicators for Family Planning Program Evaluation. Final Report of the Subcommittee on Adolescent Reproductive Health Services.* Chapel Hill, NC: The EVALUATION Project.

Thongkrajai, E., J. Stoeckel, M. Kievying, et al. 1994. *AIDS Prevention Among Adolescents: An Intervention Study in Northeast Thailand.* Women and AIDS Research Program, Research Report Series, No. 1. Washington, DC: International Center for Research on Women.

Townsend, J.W., E. Diaz de May, Y. Sepulveda, et al. 1987. Sex education and family planning services for young adults: Alternative urban strategies in Mexico. *Studies in Family Planning* 18 (2): 103–108.

U.S. Agency for International Development Center for Development Information and Evaluation. 1996. "Establishing performance targets." *Performance Monitoring and Evaluation TIPS No. 8.* Washington, DC: USAID CDIE.

U.S. Agency for International Development Center for Development Information and Evaluation. 1997. "The role of evaluation in USAID." *Performance Monitoring and Evaluation TIPS No. 11.* Washington, DC: USAID CDIE.

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

Verma, R.K., S. Sureender and M. Guruswamy. 1997. What do school children and teachers in rural Maharashtra think of AIDS and sex? *Health Transition Review* 7 (Suppl.): 481–86.

Wallerstein, N. 1992. Powerlessness, empowerment, and health: Implications for health promotion programs. *American Journal of Health Promotion* 6: 197–205.

Weiss, E., D. Whelan and G.R. Gupta. 1996. *Vulnerability and Opportunity: Adolescents and HIV/AIDS in the Developing World.* Washington, DC: International Center for Research on Women.

Weller, S. and K. Romney. 1988. *Systematic Data Collection.* Qualitative Research Methods Series 10, Newbury Park: Sage.

Wellstart International. n.d. *Tool Kit for Monitoring and Evaluating Breastfeeding Practices and Programs.* San Diego, CA: Policy and Technical Monographs.

Werner, E.E. 1994. Overcoming the odds. *Development and Behavioral Pediatrics* 15 (2): 131–36.

Wolff, J.A., L.J. Suttenfield and S.C. Binzen. 1991. *The Family Planning Manager's Handbook.* West Hartford, CT: Kumarian Press.

World Bank and the Program for Appropriate Technology in Health (PATH). March 1991. *Adult Learning and Training Techniques.* Washington, DC: World Bank.

World Health Organization. 1989. *The Reproductive Health of Adolescents: A Strategy of Action.* (Joint WHO/UNFPA/UNICEF statement.) Geneva, Switzerland: WHO.

World Health Organization. 1992. *Comprehensive School Health Education: Suggested Guidelines for Action.* (WHO/UNESCO/UNICEF consultation on strategies for implementing comprehensive school health education/promotion programmes.) Geneva, Switzerland: WHO.

World Health Organization. 1994. "Evaluation of a national AIDS programme: A methods package." *Prevention of HIV Infection.* Geneva, Switzerland: Global Programme on AIDS. WHO/GPA/TCO/SEF/94.

World Health Organization. 1996. "Programming for Adolescent Health" (Unpublished draft technical report of the WHO/UNFPA/UNICEF Study Group). Geneva, Switzerland: WHO.

World Health Organization. April 1997a. *Reproductive Health Indictors for Global Monitoring.* Report of and Interagency Technical Meeting: Division of Reproductive Health, UNDP/UNFPA/WHO/World Bank Special Programme of Research, and Development and Research Training in Human Reproduction (HRP). Geneva, Switzerland: WHO. WHO/RHT/97.27.

World Health Organization. April 1997b. *Selecting Reproductive Health Indicators: A Guide for District Managers. Field-Testing Version.* Geneva: Division of Reproductive Health, UNDP/UNFPA/WHO/World Bank Special Programme of Research, and Development and Research Training in Human Reproduction (HRP). Geneva, Switzerland: WHO/RHT/97.25.

World Health Organization. 1997. *Regional Plan of Action on Adolescent Health.* Washington, DC: Pan American Health Organization, Division of Health Promotion and Protection, Family Health and Population Program, Adolescence.

World Health Organization. 1998. *Progress in Human Reproduction Research.* Division of Reproductive Health, UNDP/UNFPA/WHO/World Bank Special Programme of Research, and Development and Research Training in Human Reproduction (HRP). Geneva: WHO.

Yoder, P.S., R. Hornick and B.C. Chirwa. 1996. Evaluating the program effects of a radio drama about AIDS in Zambia. *Studies in Family Planning* 27 (4): 188–203.

Zabin, L.S. and M.B. Hirsch. 1987. *Evaluation of Pregnancy Prevention Programs in the School Context.* Lexington, MA: Lexington Books.

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

PHOTO: Cathryn Wilcox/JHU/CCP

Appendix 1:

SAMPLING SCHEMES FOR CORE DATA COLLECTION STRATEGIES

This Appendix provides more detailed information on sampling schemes for commonly used data collection strategies in youth program evaluations. The following topics are covered:

- How to choose a systematic sample of clusters
- Cluster sampling for household surveys
- Alternative methods for choosing sample households, youth and parents
- Cluster sampling for school-based surveys
- How to allocate a proportional sample of students to schools
- Cluster sampling for health facility surveys
- Alternative methods for sampling service transactions and clients for exit interviews
- Sampling for peer education program evaluations
- Sampling for client follow-up surveys
- Sampling for focus groups and other "small group" data collection efforts
- Sampling for in-depth interviews

How to Choose a Systematic Sample of Clusters

Later sections of this Appendix will describe the use of cluster sampling in surveys of households, schools and health facilities. In cluster surveys, sample clusters are usually chosen using systematic sampling. Systematic sampling can be done with either *probability-proportional-to-size*, or *equal probability*.

Sampling with probability-proportional-to-size

Probability-proportional-to-size means that larger clusters are given a greater chance of selection than smaller clusters. The purpose of this method is to reduce the chance that study subjects associated with smaller clusters will be overrepresented in a sample. This method should be used when:

- the clusters (geographic areas, schools, etc.) vary significantly in size (for example, some clusters are at least two or three times larger than others), and
- you have information on the size of the clusters (in terms of numbers of elements) before beginning sample selection. This information is commonly called a *measure of size*.

Steps for selecting a systematic random sample of clusters with probability-proportional-to-size (PPS)

As you read the steps below for selecting a systematic random sample of clusters with probability-proportional-to-size, refer to Table 1 for an illustrative example.

(1) Prepare a list of first-stage clusters (e.g., schools, geographic areas, facilities) with a corresponding measure of size for each (**Column A**).

In the example in Table 1, there are 170 clusters to sample from. The measure of size used is the number of households in each cluster.

(2) Starting at the top of the list and moving down, calculate the cumulative measure of size. Enter these figures in a column next to the measure of size column (**Column B**). The last number in this column will be the total cumulative measure of size for all the clusters in the sampling frame. This total will be called **M**.

The total cumulative measure of size in the example in Table 1 is M = 17,219.

(3) Calculate the sampling interval (**SI**) by dividing the total cumulative measure of size (**M**) by the number of clusters to be selected (**a**).[1] That is, **SI = M/a**.

Refer to the figures from Table 1 as an example:
Total cumulative measure of size: M = 17,219
Planned number of sample clusters: a = 40
Sampling interval: SI = 17,219/40 = 430.475

(4) Select a random number (**RN**) between 1 and (**SI**). Compare this number with the measure of size column (**Column A**). The first sample cluster will be the one within whose cumulated measure of size the number (**RN**) falls.

*In the example in Table 1, the random number (**RN**) chosen between 1 and 430.475 (**SI**) is 73. The number 73 falls within the cumulated measure for Cluster No. 001 (see Columns B and C). Therefore, Cluster No. 001 is the first sample cluster (see Column D).*

(5) Choose subsequent clusters by adding progressive multiples of the sampling interval (**SI**) to the random number (**RN**) identified in Step 4; that is, RN + SI, RN + 2SI, RN + 3SI, etc. As in Step 4, select the clusters within whose cumulated measure of size these new numbers fall.

Note: When selecting sample clusters, it is important that the decimal points in the sampling interval (**SI**) be retained. When the decimal part of the sample selection number is less than .5, choose the lower number. When the decimal part of the sample selection number is .5 or greater, choose the higher number.

[1] Determining the number of clusters to sample (referred to as "a" above) depends mostly on your resource availability for data collection. The default minimum number of clusters to sample is 30, but the validity of your results will improve if you are able to sample more than 30 clusters. The more clusters you sample from, the fewer elements per cluster will need to be sampled in order to achieve your overall desired sample size, resulting in a less-biased sample and smaller design effects. If you are not familiar with these issues, it is advised that you seek outside help in determining an appropriate and feasible number of clusters to collect data from.

In the example in Table 1, RN = 73 and SI = 430.475
RN + SI = 73 + 430.475 = 503.475 (rounded down) = 503
RN + 2SI = 73 + 860.95 = 933.95 (rounded up) = 934
(see Column C)

Therefore, the clusters selected (those shown in Table 1) are Clusters 001, 005, and 009 (see Column D).

(6) Follow the procedure in Step 5 until the list of clusters has been exhausted.

Table 1
Illustrative Example of Systematic Random Sampling of Clusters with Probability-Proportional-to-Size

	A	B	C	D
Cluster Number	**Measure of Size— Number of Households**	**Cumulative Size**	**Sample Selection No.**	**Cluster Selected**
001	120	120	73 (**RN**)	X
002	105	225		
003	132	357		
004	96	453		
005	110	563	503 (**RN** + **SI**)	X
006	102	665		
007	165	830		
008	98	928		
009	115	1,043	934 (**RN** + **2SI**)	X
.	.	.		
.	.	.		
.	.	.		
170 (last)	196	17,219 (**M**)		

Sampling with equal probability
Equal probability means that all clusters are given an equal chance of selection. Sampling with equal probability should be used when:

- ➤ the clusters to be used are approximately equal in size, or
- ➤ you do not have information on the size of the clusters prior to sample selection. In other words, there is no information on the *measure of size*.

Steps for selecting a systematic random sample of clusters with equal probability
As you read the steps below for selecting a systematic random sample with equal probability, refer to Table 2 for an illustrative example.

(1) Prepare a numbered list of sites or clusters, preferably ordered geographically (e.g., by areas of a city) (**Column A**).

(2) Calculate the sampling interval (**SI**) by dividing the total number of clusters in the target group (**M**) by the number of clusters to be selected (**a**). That is, SI = M/a.

*In the example in Table 2, the total number of clusters in the target group (**M**) is 170. The planned number of clusters to be sampled (**a**) is 40. Therefore, the sampling interval (**SI**) = 170/40 = 4.25.*

(3) Select a random number (**RN**) between 1 and (**SI**). The cluster on the numbered list corresponding to this random number will be the first sample cluster.

*In the example in Table 2, the random number (**RN**) chosen between 1 and 4.25 (**SI**) is 2. Therefore, the first sample cluster is Cluster No. 002.*

(4) Choose subsequent clusters by adding progressive multiples of the sampling interval (**SI**) to the random number (**RN**) identified in Step 3; that is, RN + SI, RN + 2SI, RN + 3SI, etc.

In the example below, RN = 2 and SI = 4.25

RN + SI = 2 + 4.25 = 6.25 (rounded down) = 6
RN + 2SI = 2 + 8.50 = 10.50 (rounded up) = 11

Therefore, the clusters selected (those shown on Table 2) are Clusters 002, 006, and 011 (see Column B).

(5) Follow the procedure in Step 4 until the list of clusters has been exhausted.

Table 2
Illustrative Example of a Systematic Random Sampling of Clusters with Equal Probability

Column A	Column B
Cluster Number	Sample Selection
001	
002	X
003	
004	
005	
006	X
007	
008	
009	
010	
011	X
.	
.	
.	
M = 170 (last)	

Cluster Sampling for Household Surveys

Many outcome indicators for youth-serving programs are measured at the level of the general population of young adults, often through household surveys. Samples of parents, communities and community leaders can also be selected from the sample clusters chosen for a household survey. This method reduces costs because all data collection can be done in one visit to each sample cluster/community.

Steps involved in cluster sampling

Step 1: Define the cluster
Geographic areas with fixed boundaries (e.g., census enumeration areas, city blocks, villages) are the most commonly used clusters for household surveys.

Step 2: Develop the sampling frame
A good source for your sampling frame is the national statistics office, as most have developed lists of geographic units (census enumeration areas) for use in census taking, as well as maps and population estimates. If for some reason official census enumeration areas cannot be used, you will have to use some other list of geographic areas covered by the program, or assemble a list yourself.

Step 3: Determine how many clusters to sample
Household surveys should cover as many clusters as resources will permit—a sample of more clusters of smaller size is preferable to one with fewer clusters of larger size. In general, at least 30 clusters should be chosen. Refer to Appendix 2 for further guidance on deciding on how many clusters to choose.

Step 4: Select sample clusters
Household surveys usually involve the following two stages of sample selection:

- Selecting sample clusters
- Selecting households from each sample cluster

First stage: Clusters are chosen using *systematic sampling*, usually with *probability-proportional-to-size*, meaning that larger clusters are given a greater chance of selection than smaller clusters. Probability-proportional-to-size is used when:

- the clusters vary significantly in size (for example, some clusters are at least two or three times larger than others), and
- you have information on the size of the clusters, usually referred to as a *measure of size*.

A *measure of size* is a count or estimate of the number of elements (e.g., number of households, estimated total population, estimated number of youth) associated with each cluster. Exact counts are not necessary for use as measures of size—rough approximations are sufficient. If measures of size are not available, sample clusters may be chosen using *systematic sampling with equal probability*.

Second stage: Choose an *equal number of households* from each sample cluster, using either simple random or systematic sampling. The number of households to be chosen per sample cluster will be determined by:

- the target sample size for the survey (discussed in Chapter 6),
- the number of clusters to be sampled, and
- the expected sample "take" of youth per household.

Sampling youth

Youth can be sampled within sample clusters using a variety of methods:

- *Listing method:* Prepare a list of households in each sample cluster and choose a random or systematic sample. This method is preferable if you are able to develop a complete list of households in sample clusters.
- *Segmentation method:* Divide the sample cluster into "segments" of equal size, choose one segment at random and interview all youth found in the sample segment.
- *Random walk method:* Choose a household in the sample cluster at random as the starting point. After interviewing all youth in the household, choose the nearest household as the second sample household. Continue this process until you have interviewed the target number of youth.

The "Alternative Methods for Choosing Sample Households, Youth and Parents" section on page 250 provides a detailed explanation of how to use the above approaches.

Sampling parents of youth

To sample parents of youth, follow the same steps as outlined above—from "Define the cluster" to "Select sample clusters." Once you have chosen a sample of clusters, you can then select samples of both youth and adults within these clusters. There are two possible ways of doing this:

- Choose a sample of households and interview both youth and parents in the same sample households (less costly, and preferable if privacy can be assured).
- Choose separate samples of households in which to interview youth and parents respectively.

Sampling community leaders

Community leaders can be sampled either independently of or in conjunction with household survey efforts. If you decide to sample community leaders *separately*, you should develop a list of all community leaders and choose either a simple random or systematic sample from that list. If your sampling of community leaders is *linked* to the household survey sampling, you should select a sample of clusters for the household survey component (following the steps above) and either:

- include all community leaders from the sampled clusters, or
- take a simple random sample of community leaders in each sample cluster/community.

An advantage of the linked strategy is that it enables you to relate indicators for community leaders to indicators for youth or parents at the cluster or community level during data analysis.

Sampling to measure community-level indicators

Community-level indicators (e.g., levels of community support or mobilization for a program) may be measured through:

- general population surveys,
- interviews with key informants or community leaders, and
- observations of community activities.

As when sampling community leaders, you can sample communities located within the geographic scope of a program being evaluated, either:

- independently of other evaluation activities, or
- in conjunction with ongoing household surveys.

The advantage of the second strategy is the same as with sampling community leaders (above).

Figure 1–1
Illustrative Application of Cluster Sampling to a Household Survey of Youth

In Romania in 1996, a Young Adult Reproductive Health Survey was conducted, which required that information be gathered from independent samples of male and female youth. A two-stage cluster sample design was used.

First stage of sample selection: Choosing the clusters

- One hundred fifty-four Census Sectors were chosen as the sample clusters, using a systematic random selection procedure with probability-proportional-to-size. The measure of size used was the number of households recorded for each sector in the 1992 Census. Only female respondents were to be chosen from these sample sectors.
- Male respondents were to be selected from a separate 154 Census Sectors, which were chosen by taking the next sector in the sampling frame. Therefore, the total number of sectors chosen was 308.

Second stage of sample selection: Choosing sample elements (male and female respondents)

- It was estimated that 5.4 households would have to be contacted in order to obtain one completed interview. This estimation was reached based on 1992 Census data on the proportion of households containing one or more persons 15–24 years of age, and assuming a response rate of 90 percent.
- After the list of households in the 1992 Census was updated, a systematic random sample of households was chosen from each cluster. All eligible respondents found in sample households were included in the sample.

Comments

- The sample size for the survey in this example is much larger than would be needed for most program evaluations.
- In this example, sample households were chosen after completing an updated list of households within sample clusters (listing method). However, either the segmentation or the modified random walk method could have been used to reach the same number of households (i.e., around 80 households per sector).

Source: Serbanescu and Morris, 1997.

Alternative Methods for Choosing Sample Households, Youth and Parents

As mentioned in the previous section, the "preferred" procedure for choosing a sample of youth or households in household surveys is the *listing method*. In this method, you can use simple random or systematic sampling to choose a sample of persons or households from a complete list of all persons/households in each sample cluster.

Creating complete lists of households in sample clusters can be time-consuming, however. Therefore, this section describes two recommended alternatives: the *segmentation method* and the *modified random walk method.*

Segmentation method

In the *segmentation method*, you divide sample clusters into smaller "segments" of approximately equal size and then choose one of these segments at random from each cluster. All youth found in the households in the chosen segment are then interviewed.

The advantages of this approach are:

- ➤ you do not have to create a list of all households in each cluster, and
- ➤ a sample is produced in which all respondents have an equal probability of selection (that is, a self-weighting sample).

Steps for choosing sample persons/households using the segmentation method

(1) **Calculate the number of segments to be created.** Divide the number of households recorded in the last census by the target segment size. The target segment size is the number of households you anticipate having to contact in order to find the desired number of youth in each cluster. The resulting number will be the number of segments to be created. For example, if the last census indicated that there were 250 households in the cluster and the target segment size was 40 households, then you would need to create 6 segments. (Note: In performing this calculation, decimal numbers of segments should be rounded to the nearest whole number.)

(2) **Update the cluster map.** Using a map of the cluster, verify/update the external boundaries of the cluster and add any internal features that may be useful for dividing the cluster into easily recognizable segments (e.g., roads, streams, etc.)

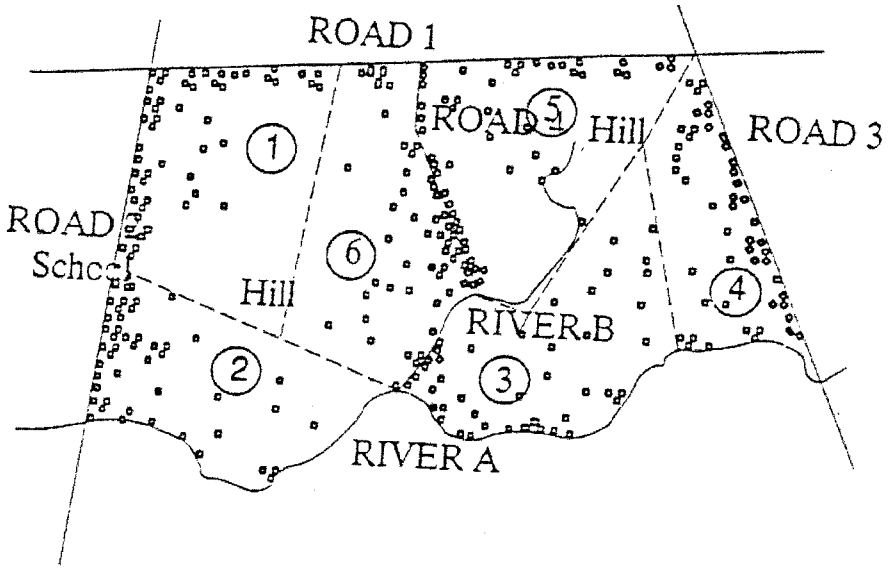
(3) **Count the number of households and indicate their location in the cluster on the map.** This allows you to divide the cluster into segments with approximately equal numbers of households.

(4) **Based on the cluster map, divide the cluster into equal-sized segments.** The number of segments to be used is the number determined in Step 1 above.

(5) **Choose one segment at random.**

(6) **Interview all households located within the boundaries of the randomly chosen segment.** Note: It is possible that the segments formed using the segmentation method might correspond to sub-groupings of the population in terms of economic status, religion, ethnicity, etc. While there is a danger that this may bias the survey results, it is to be anticipated that the various sub-groups will be appropriately "represented" over a sample of 30 or more segments.

Figure 1–2
Example of a Hypothetical Cluster that Has Been Divided Into Six Segments



Source: UNICEF. 1995. *Monitoring Progress Toward the Goals of the World Summit for Children: A Practical Handbook for Multiple-Indicator Surveys*. New York: UNICEF.

Modified random walk method

In the *modified random walk method*, sample persons/households are selected by randomly choosing a starting point, interviewing all youth/parents in the nearest household and then moving on to neighboring households until you have reached the target sample number. A map of a sample cluster is used to indicate a number of possible starting points at various, easily identifiable locations in the cluster (see Figure 1–3 for an illustrative example). This method may be familiar to many readers as the sample selection method used in Expanded Programme on Immunization (EPI) Cluster Surveys.

The primary advantages of this variant over the more commonly used EPI sampling procedure are:

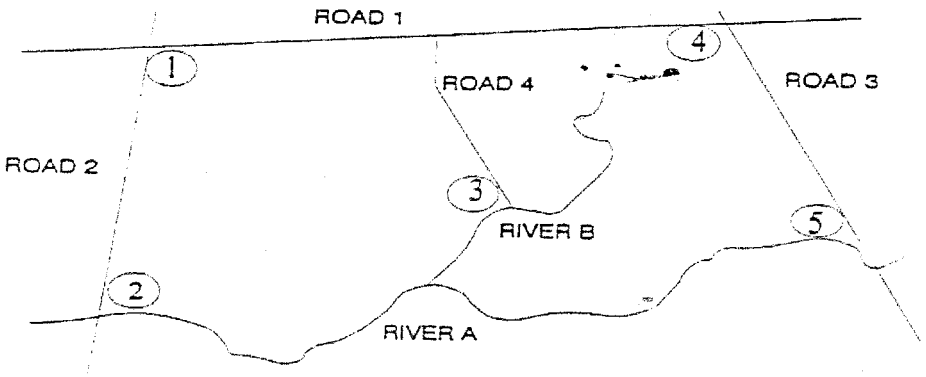
- the selection of starting points is truly random (and can be performed in advance of survey field work by supervisory personnel), and
- it does not always start in the center of the cluster, a practice often followed in EPI cluster surveys.

Note: The modified random walk method may be further strengthened by randomly choosing two or more starting points, and dividing the target sample size equally among these points.

Steps for choosing sample persons/households using the modified random walk method

- ➤ Randomly choose a starting point and direction of travel within a sample cluster.
- ➤ Choose the nearest household to the starting point as the first sample household.
- ➤ After completing the interview(s) in this first household, choose the household whose front door is nearest the first sample household as the second sample household.
- ➤ After each interview, continue to choose the next nearest household until the target number of youth/parents has been identified.

Figure 1–3
Map of Hypothetical Sample Cluster Showing Possible Starting Points



Source: UNICEF. 1995. *Monitoring Progress Toward the Goals of the World Summit for Children: A Practical Handbook for Multiple-Indicator Surveys*. New York: UNICEF.

Cluster Sampling for School-based Surveys

Several data collection protocols may be used to measure program design, system functioning and program implementation indicators for school-based programs, such as:

- ➤ curricula reviews,
- ➤ reviews of program records/statistics,
- ➤ facility assessment questionnaires or checklists, and
- ➤ surveys of students, teachers and administrators.

School surveys of students are typically used for measuring program outcome indicators.

The following sections describe the procedures for selecting a sample of schools and sample elements within them (e.g., students, teachers, teacher-student transactions, etc.).

Selecting a sample of schools

Step 1: Define the cluster
Schools (or *classes*) are the most commonly used clusters for school surveys.

Step 2: Develop the sampling frame
The sampling frame will be a list of all schools covered by the program.

Step 3: Determine how many schools (clusters) to sample
If you have decided that you need to sample schools, you will then need to determine how many schools to include in your sample:

- In large programs, 30 or more schools should be included in the sample.
- In programs covering fewer than 30 schools, as many as possible should be included in the sample.

If you are going to sample *fewer than 20 schools*, there will be little advantage in selecting schools randomly. In this case, a sample of schools that is "representative" with regard to factors such as size, location, socioeconomic level of the student body, etc. will generally be just as valid as a random sample.

Step 4: Select sample clusters
If it is not possible to include all schools in the evaluation (for example, in the case of national or regional FLE programs), you will need to limit the evaluation to a sample of schools. How many stages of sampling you will need depends on:

- the number of schools covered by the program,
- the size of schools in terms of numbers of students and classes, and
- the resources available for the program evaluation.

A cluster sampling scheme similar to that used in household surveys is the most common sampling method for school surveys. You should select schools using systematic sampling with *probability-proportional-to-size*, using data on school enrollment as your *measure of size*. (Refer to "Cluster Sampling for Household Surveys" section on page 247 for details.)

Sampling students

Step 1: Determine how many students to select from each school
If you have decided to sample students within sample schools, you will then need to determine how many students from each school to include in your sample. In school surveys, you should *not* select an equal number of students per sample school. Instead, the sample size should be *proportional* to the number of students in the school. In other words, larger schools should receive a larger sample allocation and smaller schools a smaller allocation. This is because the purpose of program evaluations is to measure program performance and outcomes for program schools as a whole, not per individual school. See the next section on page 255 for a detailed explanation of how to allocate a sample of students to schools.

Step 2: Select a sample of students in each sample school

If you are planning to use self-administered questionnaires, it may be possible to include all students in the sample schools. If the sample schools are too large to collect and process data from every student, you will need to sample the students through one of the following methods:

- Select a simple random or systematic random sample from a list of students at each sample school, or
- Choose a simple random or systematic random sample of classes from each school, and include *all* students in these sample classes. This option is simpler, especially if self-administered questionnaires are to be used.

The recommended procedure for sampling classes (Option 2) is as follows:

- Create a list of the classes in each sample school, ordered by grade.
- Then, assuming that classes within schools are of approximately equal size, use systematic sampling with equal probability to select a sample of classes.

If you choose Option 2, the number of classes or sections to be chosen in a given school will be determined by:

- the target sample size for the school, and
- the size of classes in the school.

For example, if a school has been allocated a sample of 100 students, and each class at the school averages 25 students, then four (4) classes should be chosen for the survey. An illustrative example of this procedure is provided in Section 1.5.

Sampling teachers or school administrators

Indicators related to teacher knowledge, attitudes, perceptions and performance in implementing a program may be measured through self-administered questionnaires or personal interviews with teachers.

Selecting a sample of teachers

If the number of schools in the program being evaluated is small, you can include all teachers at sample schools in teacher surveys. For larger-scale programs, the two possible ways to select a sample of teachers are:

- using simple random or systematic random selection to select teachers from a list of all teachers at each sample school, or
- choosing teachers associated with the sample classes/sections selected for student surveys. One advantage of this strategy is that indicators for students and teachers can be directly related to one another during analysis.

When measuring indicators related to *administrators*, all administrators in sample schools should be included, as their number per school is usually quite small.

Sampling parents of students

If you are already planning to conduct household surveys of youth (and if more than 50 percent of school-aged youth in the general population have been exposed to the program being evaluated), one possibility is to interview parents of school-attending youth found in sample households. Two other, more efficient approaches are to:

- ➤ select a simple random or systematic random sample of students from a list of students at program schools (or a sample of schools), and interview the parents of these students, or to
- ➤ choose a simple random or systematic random sample of classes from program schools, and interview parents of *all* students in the sample classes.

As with other stakeholders, you have the option of sampling parents independently of or in conjunction with your samples of students. The advantage of the latter strategy is that it allows you to relate indicators measured for students with those for their parents. Figure 1–4 provides an illustrative application of some of the sampling strategies described above for an evaluation of a school-based FLE program.

Figure 1–4
Illustrative Example of an "In-Class" School Survey of Youth

As part of an evaluation of the national sex education/family life education program recently implemented in public schools in Peru, a KAP-type survey was conducted in a sample of public-sector secondary schools.

Selecting sample schools (clusters): A total of 34 schools—10 from Metro Lima and 24 from cities in the interior of the country—were selected, using a systematic random selection procedure with probability-proportional-to-size.

Selecting sample classes/sections within each school: In each sample school, one class/section from each of the five grades at the secondary level of the education system in Peru was chosen to be included in the survey, using simple random sampling.

Total sample size: The estimated average class size was 30 students. Therefore, the expected sample size per school was n = 150 (5 classes x 30 students per class). The total expected sample size was n = 5,100 students (34 schools x 150 students per school).

Self-administered questionnaires were used in the survey.

Source: Ministry of Education, Peru, and FOCUS on Young Adults Program, 1998.

How to Allocate a Proportional Sample of Students to Schools

As mentioned in the "Cluster Sampling for School-based Survey" section, the target sample size of students from each sample school should be proportional to the size of the school, rather than sampling an equal number of students from each school. This means that a larger sample of students will be taken from large schools and a smaller sample from small schools, which will allow you to measure program performance and outcomes for program schools as a whole, rather than per individual school.

Table 3
Illustrative Example of Proportional Sample Allocation to Schools

A	B	C	D
School	**Number of Students Enrolled**	**Proportion of Students in Sample School**	**Target Sample Size (Sample Size = 2,000)**
001	296	.015	30
002	845	.043	86
003	692	.035	70
004	399	.020	40
005	908	.046	93
.	.	.	
.	.	.	
.	.	.	
030	773	.040	79
Total	19,565	1	2,000

Steps for allocating a proportional sample of students to schools
As you read the steps below for allocating a proportional sample of students to schools, refer to the illustrative example in Table 3.

(1) Prepare a list of sample schools (**Column A**), preferably ordered geographically (e.g., by areas of a city), showing the number of students enrolled at each school (**Column B**).

(2) For each sample school, calculate the proportion of the total number of students in all sample schools enrolled in that school (**Column C**). This number is found by dividing the number of students in each particular school by the total number of students in all the schools combined (**Total of Column B**).

For example, take School 002 from Table 3:

845 (number of students in School 002) divided by 19,565 (total number of students) = .043 (the proportion of total number of students in School 002)

(3) Multiply the proportion for each school obtained in Step 2 (**Column C**) by the target sample size for the survey. The result will be the target sample size for each school (**Column D**).

Again take School 002 as an example:

.043 (proportion of students in School 002) times 2,000 (target sample size for the survey) = 86 (target sample size for School 002)

Cluster Sampling for Health Facility Surveys

For health facility-based programs, the measurement of most program design, system functioning and implementation indicators will require visits to the program facilities. The data collection protocols that may be used during such visits include:

- ➤ reviews of program records/statistics,
- ➤ facility assessment questionnaires or checklists,
- ➤ interviews with service providers and administrators,
- ➤ observations of transactions between service providers and youth clients, and
- ➤ interviews with clients.

Unlike school programs, follow-up surveys of clients will normally be required to measure program outcome indicators, since client contact with program facilities is for a much briefer period of time. Such surveys are discussed in a separate section later in this chapter.

Sampling facilities

Step 1: Define the cluster
Health facilities are the most common cluster for health facility surveys and assessments.

Step 2: Develop the sampling frame
The sampling frame will be a list of all health facilities covered by the program.

Step 3: Determine how many facilities (clusters) to sample
As in the case of household and school surveys, as many facilities as possible should be included in the sample. In large programs, 30 or more facilities would ideally be covered. In programs with fewer than 20 facilities, a sample of facilities that is "representative" of all facilities in terms of size, location, socioeconomic level of the catchment area served, etc. will suffice.

Step 4: Select sample clusters
For small-scale programs, it may be possible to include all facilities in the sample. For larger-scale programs covering multiple facilities, it may be necessary to limit the evaluation to a sample of facilities. A two-stage cluster sampling scheme similar to that used in school surveys is the most common sampling method for health facility surveys. You should select facilities using systematic sampling with *probability-proportional-to-size*, using the volume of youth clients at health facilities as your *measure of size*. (Refer to the "Cluster Sampling for Household Surveys" section on page 247 for details.) If information on client volume is not available, facilities should be chosen using either simple random or systematic random sampling (with equal probability).

Sampling service providers

Indicators related to service provider knowledge, attitudes, program implementation practices and perceptions about the program may be measured through self-administered questionnaires or personal interviews with service providers.

Selecting a sample of service providers

If the number of program facilities and personnel involved is small, you should include all service providers who serve youth at sample facilities.

In larger-scale programs, you can choose a sample of service providers, either:

- by using simple random or systematic random selection to choose providers from lists of service providers at each sample facility, or
- by interviewing all service providers who happen to be present on a randomly chosen day that a sample facility is visited.

The latter strategy is often used when *several data collection protocols are to be used* (e.g., facility assessment checklists, interviews with service providers and youth clients and observations of service transactions), as all may be conveniently administered during a one- or two-day visit to each sample facility.

Sampling service transactions

The most common method of observing service transactions is to deploy trained observers at sample facilities. The *observation period* can either be:

- a fixed time interval (e.g., one or two days per facility), or
- the length of time necessary to observe a pre-determined number of service transactions (e.g., 10 transactions per facility).

Selecting observation days

In either case, you should randomly choose the day(s) of the week on which observations are made at each sample facility to ensure that observations are spread out over different staff. The challenge of this measurement approach is that the data collection schedule depends entirely on when and how often youth come to the clinic for services. If youth clients visit the facility infrequently, using the *fixed time interval* may result in relatively few transactions being observed, while the *service transaction quota approach* may require several days of observation for quotas to be reached. The use of these observation protocols is made somewhat easier when facilities have special hours for youth clients, as service transactions with these clients are concentrated into shorter time intervals.

Sampling for mystery client observations

In the *mystery client observation approach*, youth are recruited to visit health facilities in the role of clients, after which information on what transpired during service contacts is recorded. Mystery client observations are preferable to the observation approach described above when:

- it may be difficult to observe sufficient numbers of service transactions with youth clients during visits on randomly chosen days, and/or
- the observation of service transactions by third parties may influence what happens during the transaction.

Selecting mystery client observation days

Once a sample of facilities has been chosen, randomly choose days of the week and times of day for mystery client visits to be made. Mystery clients should be asked to play out a variety of scenarios, so that different types of services can be observed. Note, however, that the youth recruited as simulated clients should not have to undergo invasive procedures at health facilities (e.g., pelvic examinations or blood tests). To avoid this risk, mystery client observations should be used to assess only limited types of services.

Figure 1–5
Illustrative Application of a Sampling Strategy for Evaluating a Health Facility-Based Program

In 1998, the Moroccan Ministry of Health conducted an assessment of the quality of services being provided to both youth and older adults at MOH facilities.

Methods of data collection: Data were gathered using facility inventory checklists, observations of service transactions and exit interviews with clients.

Selecting a sample of facilities: A sample of facilities was chosen by first choosing a sample of six provinces, and then choosing a sample of facilities within these provinces. Sample facilities were chosen with probability-proportional-to-size, using the estimated daily client volume for reproductive health services as the measure of size. A total of 96 sample facilities were chosen.

Collecting data in each sample facility: Each sample facility was visited for one randomly chosen day, during which a team of three persons undertook the various data collection tasks. The number of teams assigned to each facility was based on the expected client volume, with high-volume facilities being assigned more than one team. All transactions involving reproductive health services conducted on the day of the visit were observed, and all clients were interviewed as they left the facility.

Sampling for client exit interviews

Another way to measure program implementation- and client response-type indicators is through interviews conducted with youth as they leave health facilities after having received services. The primary sampling objective is to obtain feedback on a random sample of service transactions. This is typically accomplished by choosing random samples of facilities and days of the week for conducting interviews with clients. As was the case for direct observation of service transactions, however, the time required to obtain a target number of interviews depends entirely on when and how often youth come to the facility to receive services.

Figure 1–5 provides an illustrative application of a sampling strategy for evaluating a health facility-based program.

Alternative Methods for Sampling Service Transactions and Clients for Exit Interviews

When conducting observations of client service transactions or client exit interviews in facility surveys, it is not always feasible to create a list of subjects from which to choose a sample. In these situations, you can use either *quota sampling* or a *take-all strategy* as alternative approaches for the second stage of sample selection.

Quota sampling

In *quota sampling*, service transactions are observed or clients are interviewed at sample facilities until a pre-determined quota of observations or exit interviews has been reached. These interviews or observations are conducted at randomly chosen points in time (e.g., days). This strategy will result in a fixed number of observations per facility, but the length of time required to reach the quota will vary across facilities, depending on how often and how many youth come to the facility for services.

The take-all strategy

In the *take-all strategy*, all clients who happen to appear at a particular facility on a randomly chosen day are included in the sample (irrespective of how many clients there are). This strategy will result in a variable number of sample elements per cluster. In order for this approach to work, you will need reasonably accurate information on the typical or average number of sample elements associated with the sites or clusters for a particular target group. This information is needed in order to ensure that, on the one hand, a sufficient number of clusters have been chosen for the survey, while, on the other hand, the overall target sample size is not grossly exceeded.

Figure 1–6
Summary of Alternative Second-Stage Sample Selection Procedures

Sampling Approach	Advantages	Disadvantages
Quota sampling	➤Tight control of sample size ➤No need for a list of sample elements	➤May require multiple visits to a site to reach sample size quota ➤Higher danger of bias
Take-all approach	➤Number of visits to sample sites limited to one ➤No need for a list of sample elements	➤No control over sample size ➤Higher danger of bias

Sampling for Peer Education Program Evaluations

Process evaluation efforts for peer education programs are usually intended to assess how well peer promoters are prepared (i.e., their recruitment, training and supervision), as well as peer promoters' own knowledge, attitudes, communication skills, etc. The methods often used to conduct this assessment are:

- interviewing random samples of peer educators (to gather the bulk of the information in process evaluation efforts),
- referring to records maintained by peer educators, or conducting surveys of youth in the appropriate target audience (to assess outputs of peer education efforts, in terms of numbers and characteristics of peers reached), and
- conducting survey interviews with clients (to assess outcomes of peer promotion activities).

Peer education interventions differ from other types of youth-serving programs in that:

- the setting or medium through which "contacts" are made between peer educators and clients varies across different types of program models (e.g., schools, community settings, youth organizations or facilities, outreach activities or the mass media);
- contacts with clients are often made on the basis of opportunity, as opposed to regular service schedules or training plans; and
- in some forms of peer education (e.g., through the mass media), there is no direct contact between peer educators and clients.

Sampling peer educators and their clients

The primary sampling concern for evaluating peer education program efforts is to ensure that "representative" samples of peer educators and clients are selected. You can choose a *sample of peer educators* by doing *simple random sampling* from a list of peer educators associated with the program being evaluated. If you want to assess the performance of different categories of peer educators, you can take simple random samples of peer educators in each of the different categories.

The preferred strategy for *sampling clients* (in order to measure outcome indicators) will depend on both the type of peer education program and the coverage/reach of the program. There are several options:

- *For mass media-based programs,* general population surveys of youth are the preferred means of measuring program outcomes. The relevant sampling strategies are the same as those described earlier for conducting household surveys.
- *If the program reaches more than 50 percent of the intended target audience,* target audience surveys will usually be the most efficient method of evaluation. Such general audience surveys will be more feasible in programs implemented in schools, youth organizations, workplace-based programs and other settings where the target audience congregates in one location (at least occasionally).

➤ *If the expected program coverage level is moderate or low (less than 50 percent),* or *for programs where clients do not congregate in fixed locations* (e.g., community-based programs), client follow-up surveys are the only way to measure program outcomes. Sampling strategies for such surveys are discussed in the next section.

Sampling for Client Follow-up Surveys

If you want to measure medium- and long-term outcome indicators for youth program evaluations, you need to have a mechanism for maintaining or re-establishing contact with program clients. This task is somewhat easier in school- or workplace-based programs, since at least some clients will maintain contact with schools or workplaces over time. However, a follow-up mechanism for such programs is still needed in order to be able to measure outcomes for clients who have either changed schools, graduated, dropped out of school or left a place of employment. This is especially important, as the reason for having left school or work may be related to adverse reproductive health outcomes (e.g., female students or workers experiencing an unwanted pregnancy).

In a *follow-up survey*, you will:

➤ choose a sample of program clients,

➤ establish contact with them, and

➤ conduct a survey interview. (Note: Qualitative data collection methods could also be used.)

Sampling clients for follow-up surveys

To sample clients for follow-up surveys, you can do one of two things:

➤ Use simple sampling schemes, such as simple random sampling from a list of program clients. This method will be sufficient in most cases.

➤ Choose a sample of peer educators, and then choose a sample of their contacts in a second stage of sample selection.

If you want information about specific sub-groups of clients, you can classify them on the basis of characteristics that may be important to program outcomes (e.g., age, gender, ethnic group, economic status) and take samples from each group. However, if you want to compare outcomes for sub-groups of youth, it will be necessary to ensure that an adequate sample size has been obtained for each group. This issue is addressed in greater detail in the discussion of sample size requirements in Chapter 6.

Length of the observation period

One challenge of client follow-up surveys is being able to establish contact with and successfully interview a sufficient number of clients. In order to measure long-term effects, an observation period of three to five years may be needed. However, the longer the observation period, the greater the proportion of clients that you will not be able to locate. Therefore, for practical purposes, follow-up periods should be limited to two to three years (unless a program evaluation has sufficient resources to track clients for longer periods).

Adjusting sample size to anticipate drop in follow-up rate
Since you might not be able to re-establish contact with some clients in follow-up surveys, sample sizes for follow-up surveys should be increased (usually by around 25 percent) in order to at least partially compensate for this expected loss. This increase will compensate only for the loss of sample size, but not for bias. It is therefore important to try to keep follow-up loss rates as low as possible.

Sampling for Focus Groups and Other "Small Group" Data Collection Efforts

Use non-probability sampling methods
Non-probability sampling methods are more commonly used than probability methods when choosing respondents for focus group discussions and related small-group data collection methods (e.g., pile sorts and free-listing). As mentioned in Chapter 6, assembling randomly chosen respondents in a given location to conduct focus groups or other small-group methods is difficult.

Focus groups based on sub-groups in the target population
Stratification or market segmentation is a key aspect of sampling for small-group methods. At least one focus group is formed for each of the key sub-groups in the population under study (this could pertain to different sub-groups of clients, as well sub-groups in the general population). For example, for youth-serving programs, relevant sub-groups might consist of youth of different ages, genders and economic-status categories, parents and community leaders.

Focus group participants should also be sufficiently well "spread out" regarding other characteristics that might influence discussion responses. You might want to hold focus group discussions in different parts of a geographic area covered by a program, at different times of the day, etc. and—to the extent feasible—randomly choose eligible participants. Although this will add to the cost of data collection, it will result in data that can be more reliably generalized than the data reached through convenience sampling.

Sampling for In-Depth Interviews

In-depth interviews involve a less structured approach to survey data collection, allowing respondents to answer questions in detail and in their own words. As with focus groups, respondents for in-depth interviews are often chosen using *non-probability sampling methods*. The advice concerning "spreading out" the sample across key categories of respondents in focus group surveys also applies to in-depth interviews. If you are using probability sampling for other program evaluation data collection purposes, you may have an opportunity to use more rigorous sampling methods for in-depth surveys through sub-sampling. For example, if you are conducting household, school or facility surveys using probability sampling methods, a small proportion (i.e., 5 or 10 percent) of respondents might then be chosen for in-depth interviews.

Appendix 2:

HOW TO CALCULATE SAMPLE SIZE REQUIREMENTS

In Chapter 6, a sample-size table was provided. In this table, the levels of statistical significance and power were set at .90. To compute sample size requirements for different levels of significance and power, the following formula may be used:

$$n = D\,[Z_\alpha\,(2P\,(1-P))^{1/2} + Z_\beta\,(P_1\,(1-P_1) + P_2\,(1-P_2))^{1/2}]^2 / (P_2 - P_1)^2$$

Where:

D = design effect;

Z_α = the z-score corresponding to the probability with which it is desired to be able to conclude that an observed change of size $(P_2 - P_1)$ would not have occurred by chance;

$P = (P_1 + P_2) / 2$;

Z_β = the z-score corresponding to the degree of confidence with which it is desired to be certain of detecting a change of size $(P_2 - P_1)$, if one actually occurred.

P_1 = the estimated proportion at the time of the first survey; and

P_2 = the proportion at some future date such that the quantity $(P_2 - P_1)$ is the size of the magnitude of change it is desired to be able to detect;

Standard values of Z_α and Z_β for use in the above formula are provided in Table 1. A look-up table showing sample sizes needed per survey round for different combinations of significance and power is provided in Table 2.

Table 1
Values of Z_α and Z_β

	Z_α		Z_β	
α	One-Sided Test	Two-Sided Test	β	Z_β
.90	1.282	1.645	.70	0.53
.95	1.645	1.960	.80	0.84
.975	1.960	2.240	.90	1.282
.99	2.326	2.576	.95	1.645
			.975	1.960
			.99	2.326

Table 2
Sample Size Requirements for Selected Combinations of P_1, P_2, α and β

P_1	P_2	Combinations of α and β (α/β)			
		95/90	95/80	90/90	90/80
.10	.20	432	312	331	228
.10	.25	216	156	165	114
.20	.30	636	460	485	336
.20	.35	299	216	229	158
.30	.40	773	558	594	408
.30	.45	352	255	270	186
.40	.50	841	607	646	444
.40	.55	375	271	288	198
.50	.60	841	607	646	444
.50	.65	367	266	282	194
.60	.70	773	558	594	408
.60	.75	329	238	253	174
.70	.80	636	460	485	336
.70	.85	261	189	200	138
.80	.90	432	312	331	228
.80	.95	163	118	125	86

Note: Sample sizes shown assume a design effect of 2.0.

What magnitude of change ($P_2 - P_1$) should be measured?

The quantity ($P_2 - P_1$) is the minimum change in a given indicator that it is desired to measure in successive surveys with a specified degree of certainty. As the value of ($P_2 - P_1$) decreases, the required sample size increases. Thus, for small values of ($P_2 - P_1$), the required sample size will be quite large. Accordingly, for practical reasons, measuring magnitudes of change in behavioral indicators on the order of 10 to 15 percentage points is recommended as the minimum for target group survey efforts, as attempts to measure changes of smaller magnitudes with adequate precision are likely to exceed the resources available in many—in fact, most—efforts.

It will be noted that the magnitude of change parameter specified for sample size determination purposes may or may not correspond to program targets with regard to the indicator in question. In some cases, a program might aspire to change an indicator by only a small amount. For example, where condom use is running only 5 percent in a given setting, it might be quite satisfactory to increase it to 10 percent over a two- to three-year period. Nevertheless, because the sample size required to detect a change of 5 percentage points may be larger than the available resources can support, the parameter ($P_2 - P_1$) might be set to 10 or 15 percentage points in determining sample size requirements for surveys. In this situation, even though the program target of increasing condom use by 5 percentage points may have already been reached, it will not be possible to conclude statistically that the indicator has changed until a change of 10–15 percentage points has been realized (unless, of course, additional resources can be found to support surveys with larger sample sizes).

In cases where larger changes in indicators are expected, it may be desired to increase the magnitude of change parameter in the sample size calculations, thereby decreasing the sample size needed. It should be recognized, however, that doing so will jeopardize the ability to detect smaller changes that may in fact be programmatically significant. For example, if a program aspires to increase condom use by 25 percentage points over a five-year period and accordingly sets ($P_2 - P_1$) equal to 25 percentage points, changes of 10 percentage points realized over the first two years of the program will not be measurable with statistical significance.

Note also that some programs do not have explicit targets for indicators, and thus sample size requirements will be driven primarily by resource and statistical factors. In such cases, the recommended "generic" target of 10–15 percentage points of detectable change is intended as a practical benchmark that should be within the resource levels available for data collection for most programs.

Determining starting or baseline levels of indicators (P_1)

A second issue concerns the choice of a starting value of an indicator being monitored, that is, P_1. Ideally, this choice is based on information available from other surveys conducted in the study setting. Where such information is unavailable, an informed guess must be made. In choosing a value for P_1, the recommended course of action is to err toward assigning P_1 a value of .50, because the variances of indicators measured as proportions are maximized as they approach .50. Thus, erring toward .50 provides a measure of insurance that the sample size chosen will be sufficient to satisfy the measurement objectives of the survey, even if the estimate of P_1 used is erroneous. The safest course would, of course, be to choose P_1 = .50 for all indicators. However, this would result in samples that are much larger than needed in the event that the actual value of P_1 is very different from .50. Thus, the recommended approach is to make the best guess based on available information, and err toward .50 in selecting values of P_1.

Design effects

A third issue concerns the *design effect* (D) to be used. The design effect provides a correction for the loss of sampling efficiency, resulting from the use of cluster sampling as opposed to simple random sampling. Thus, D may be simply interpreted as the factor by which the sample size for a cluster sample would have to be increased in order to produce survey estimates with the same precision as a simple random sample.

The magnitude of D depends on two factors: (1) the degree of similarity or homogeneity of elements within clusters, and (2) the number of sample elements to be taken from each cluster. The initial factor, the homogeneity of elements within clusters, is a population characteristic over which the survey taker has no control. Prior methodological research indicates that most population characteristics tend to cluster, and thus the prudent course is to assume that some degree of homogeneity within clusters exists. The second parameter, the number of elementary units chosen per cluster, is largely within the control of the survey taker and is an important consideration in the sample design for any survey (see below for further discussion).

What size design effect should be used in estimating sample sizes? Ideally, an estimate of D for the indicators of interest could be obtained from prior surveys in any given setting. Short of this, "typical" values from surveys conducted elsewhere could be used. If no information is available on the magnitude of design effects for the indicators of interest, the use of a "default" value is recommended. In many cluster surveys, a default value of D = 2.0 is used. Assuming that cluster sample sizes can be kept moderately small in target group survey applications (e.g., not more than 20–25 elements per cluster), the use of a standard value of D = 2.0 should adequately compensate for the use of cluster sampling in most cases.

Should one- or two-tailed z-score values be used?

In program evaluation situations, there is good reason to anticipate the direction in which key indicators will change. Accordingly, in the example of sample size computations in Chapter 6, one-tailed values of Z_α were used. This will result in a smaller sample size than if the corresponding two-tailed values had been used. As a general rule, one-tailed tests should be used only when there is a clear rationale for expecting a change in a given indicator in one direction, for example, when an intervention of substantial magnitude and aimed at a given target group has been implemented. Otherwise, the prudent course of action is to use two-tailed values of Z_α.

Power

A point warranting special attention in survey undertakings in which a priority objective is to measure changes in indicators over time is that of *power*. Unless sample sizes are sufficient to be able to detect changes of a specified size, the utility of repeated surveys as a monitoring tool is compromised. To illustrate, suppose we desired to be able to measure a change of 10 percentage points in the proportion of sex workers who always use a condom with their clients. We compare two pairs of hypothetical surveys taken two years apart: one with a sample size of n = 500 per survey round and the other with a sample size of n = 200 per survey round. While both surveys might indicate the expected increase of 10 percentage points, this change may not be statistically significant at a given level of significance based on the survey with a sample size of n = 200. Thus, we would be forced to conclude that there was no significant change in this behavior over the period study, when in fact there was a real increase that was simply not detectable. To ensure sufficient power, a minimum value of Z_β of .80 should be used, and .90 would be preferable if resources permit.

Appendix 3:

REFERENCE SHELF

Bertrand, J.T., R.J. Magnani and N. Rutenberg. September 1996. *Evaluating Family Planning Programs with Adaptations for Reproductive Health.* Chapel Hill, NC: The EVALUATION Project.

Brindis, C., J.J. Card, S. Niego and J.L. Peterson. 1996. *Assessing Your Community's Needs and Assets: A Collaborative Approach to Adolescent Pregnancy Prevention.* Los Altos, CA: Sociometrics Corporation.

Brindis, C., J.L. Peterson, J.J. Card and M. Eisen. 1996. *Prevention Minimum Evaluation Data Set: A Minimum Data Set for Evaluating Programs Aimed at Preventing Adolescent Pregnancy and STD/HIV/AIDS, 2nd Edition.* Los Altos, CA: Sociometrics Corporation.

Feuerstein, M.T. 1994. *Partners in Evaluation: Evaluating Development and Community Programmes with Participants.* London and Basingstoke: The MacMillan Press Ltd.

Fisher, A., J. Laing, J. Stoeckel and J. Townsend. 1991. *Handbook for Family Planning Operations Research Design, 2nd Edition.* New York: The Population Council.

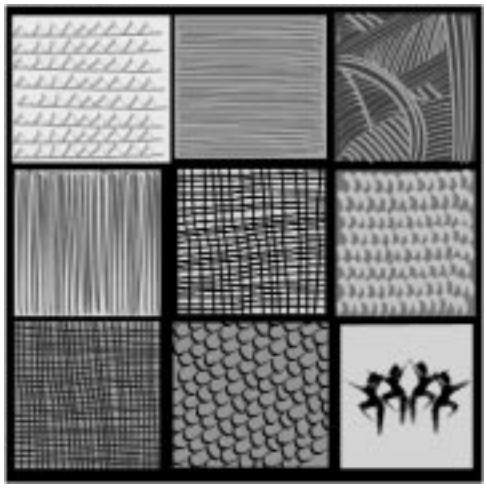
García-Núñez, J. 1992. *Improving Family Planning Evaluation.* West Hartford, CT: Kumarian Press.

Nelson, K., L. MacLaren and R. Magnani. 1999. *Assessing and Planning for Youth-Friendly Reproductive Health Services.* Washington, DC: FOCUS on Young Adults.

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

Rossi, P. and H. Freeman. 1993. *Evaluation: A Systematic Approach, 5th Edition.* Newbury Park, CA: Sage Publications.

Shah, M.K., R. Zambezi and M. Simasiku. 1999. *Listening to Young Voices: Facilitating Participatory Appraisals with Adolescents on Reproductive Health.* Washington, DC: Care International in Zambia and FOCUS on Young Adults.



Appendix 4:

EVALUATION WEB SITES

General Sites

http://www.eval.org
The American Evaluation Association, an international professional association of evaluators, is devoted to the application and exploration of program evaluation, personnel evaluation, evaluation technology and other forms of evaluation.

http://www3.sympatico.ca/gpic/gpichome.htm
This site offers links to many Web resources on evaluation, brought to you by Government Performance Information Consultants.

http://www.unitedway.org/outcomes/
The United Way's Resource Network on Outcome Measurement offers a guide to resources for measuring program outcomes for health, human service and youth- and family-serving agencies. Their manual, *Measuring Program Outcomes: A Practical Approach*, can be ordered here.

http://www.unites.uqam.ca/ces/mainpage.html
The Canadian Evaluation Association is dedicated to the advancement of evaluation for its members and the public. (This site is also available in French.)

http://hogg1.lac.utexas.edu/Gen/
The Grantmakers Evaluation Network (GEN) is an affinity group of the Council on Foundations. The purpose of GEN is to promote the development and growth of evaluation in philanthropy. GEN will seek to leverage, expand and diversify the sources of philanthropic dollars for evaluation and to build the capacity of members and others in its pursuit.

http://www.wmich.edu/evalctr/
The Evaluation Center, located at Western Michigan University, is a research and development unit that provides national and international leadership for advancing the theory and practice of evaluation, as applied to education and human services.

http://www.socio.com/
This is Sociometrics' home page. Click on "Evaluation Resources" for a description of evaluation resources available directly from Sociometrics.

http://www.stanford.edu/~davidf/empowermentevaluation.html
The American Evaluation Association has a Collaborative, Participatory and Empowerment Evaluation topical interest group that is dedicated to the exploration and refinement of collaborative, participatory and empowerment approaches to evaluation.

http://www.inetwork.org/
Innovation Network, Inc. (InnoNet), is an organization whose mission is to enable public and nonprofit organizations to better plan, execute and evaluate their structures, operations and services. InnoNet has two services to meet this end: a search service to find model programs, and an evaluation service that guides agencies through a planning and evaluation process. Descriptions of their evaluation methodologies and documents available for ordering are listed on this site.

http://trochim.human.cornell.edu/kb/conmap.htm
Bill Trochim is a faculty member at Cornell University; his work in applied social research and evaluation is described on this site. His published and unpublished papers, detailed examples of current research projects, useful tools for researchers, an extensive online textbook, a bulletin board for discussions and links to other locations on the Web that deal in applied social research methods are included.

http://www.freenet.tlh.fl.us/~polland/qbook.html
This site contains a complete manual, entitled *Essentials of Survey Research and Analysis: A Workbook for Community Researchers*, written by Ronald Jay Polland, Ph.D., 1998.

http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/start.htm
This site contains a complete manual, *User-Friendly Handbook for Mixed Method Evaluations* (August 1997), edited by Joy Frechtling and Laurie Sharp Westat, and developed with support from the National Science Foundation, Division of Research, Evaluation and Communication.

International Sites

http://www.wmich.edu/evalctr/ICCE
The International & Cross-Cultural Evaluation Topical Interest Group (I&CCE) is an organization affiliated with the American Evaluation Association. The purpose of the I&CCE is to provide evaluation professionals who are interested in cross-cultural issues with an opportunity to share their experiences with one another.

http://www.rrz.uni-koeln.de/ew-fak/Wiso/
This is the home page for the German Center of Evaluation (in German) at the University of Cologne. It includes the German translation of the Program Evaluation Standards of the American Evaluation Society.

http://www.dec.org/usaid_eval/
The U.S. Agency for International Development's Development Experience Clearinghouse (DEC) is a publication clearinghouse that contains references to USAID-funded documentation. The Center for Development Information and Evaluation (CDIE) publications from 1997 through 1998 are provided here and are arranged by CDIE publication series title.

http://www.unicef.org/reseval/
This site lists some of the monitoring and evaluation tools recently developed by UNICEF and its partners, including the *UNICEF Guide for Monitoring and Evaluation*.

Education Sites

http://ericae.net/
This site lists many education-related links for assessment and evaluation.

Mental Health Sites

http://www.vanderbilt.edu/VIPPS/CMHP/
The Center for Mental Health Policy is housed in the Vanderbilt Institute for Public Policy Studies of Vanderbilt University, and focuses on child, adolescent and family mental health services research. Their page has links to other mental health-related sites.

HIV/AIDS Sites

http://hivinsite.ucsf.edu/prevention/evaluating_programs/
Maintained by the Center for AIDS Prevention Studies (CAPS) at the University of California, San Francisco (**http://www.caps.uscf.edu/index.html**), this site provides tools to help plan, design and implement evaluations of HIV prevention programs.

http://www.themeasurementgroup.com/edc.htm
The Measurement Group, in collaboration with PROTOTYPES, has been funded by the Health Resources and Services Administration (HRSA) to provide help on evaluation and dissemination activities to 27 national demonstration programs on HIV/AIDS treatment services. This Evaluation and Dissemination Center is part of HRSA's activities to develop innovative models for treating HIV/AIDS.

PHOTO: Harvey Nelson