THE MONITORING
AND
EVALUATION
GUIDEBOOK

By Macro TAT

Table of Contents

List of Tables and Figures

MACRO
INTERNATIONAL INC.

Introduction:

The fact that you are reading this manual demonstrates that you have made a commitment to improve your monitoring and evaluation (M&E) skills. Effective program management is a combination of administrative, logistic, and M&E skills. M&E provides the feedback mechanism that you need to give you an indication of "where you are" in terms of managing for results. At USAID/Mali we have covered all of the basics in the USAID managing for results system and we have looked at each component as thoroughly as time has permitted. This manual is intended to guide you as a resource and serve as a jumping off point in your continuing journey into the field of methods, otherwise known as M&E.

In development work, monitoring and evaluation helps us to determine where we are going, and once we reach a destination, to assess if that is where we want to be. ***Monitoring*** involves the use of easy and inexpensive tools that help us plot our progress in meeting our implementation and goals and chart a course for the future. ***Evaluation*** helps us take a broader perspective to determine if our course is the best one, if we are doing the right things and doing them right.[1]

Evaluation as a Management Tool

The word evaluation is often shrouded in mystery and even fear. For many it conjures up images of tests and examinations where others pass judgement over us. These images and reactions are major stumbling blocks to using evaluation as a tool to assist us in making more informed decisions. A major need is to demystify the concept of evaluation. Although some may hope to coin a different term to reflect what we want to do, it may be best to face the fears and potential negative reactions straight on.

Monitoring and Evaluation is really a process of producing information that can then be referred to and used in a number of ways within an organization. M&E is considered as a means to an end, a means to improve programs, predict outcomes or benefits and a way to become more accountable.[2]

M&E is a means to an end:
Better programs
Greater benefits
More accountability
Why M&E? • To better understand organizational operations • To improve programs by identifying strengths and weaknesses • To plan for future programs by testing alternative strategies/approaches • To demonstrate the merits and limitations of programs • To reassure groups with interests at stake by building quality control safeguards

---

[1] *Child Survival,* BASICS Volume no. 5, Spring 1998 p 2.
[2] *Check Points: Evaluation as a Management Tool,* The United Way Michigan and ADS@MSU.

One word of caution about using evaluation as a management tool, in the context of Managing for Results, is that the person conducting the evaluation or writing a report should always honestly state USAID's contribution. Evaluation can easily be used for political ends to support decisions, instead of as an organizational learning tool. Using evaluation to support political interests is a symbolic action and reviewers will realize this and question the validity of your results. Also, political interests are usually short term and are bad for long term planning; especially when budgeting scarce resources that are to be divided up between programs and evaluation activities.

In this manual the various skills transfer modules are designed to take you through the Monitoring and Evaluation process to help you better Manage for Results at USAID. Each module is a separate unit therefore; you may hop from one module to another without following an ordered sequence.

- Module One is an introduction to Monitoring and Evaluation the basic concepts, terms and tools used

- Module Two is based on the logic of using a Results Framework and selecting appropriate indicators

- Module Three A Framework For Planning Project Data Needs

- Module Four Project Monitoring for Activity Managers

- Module Five Using Case Studies to Illuminate Results

- Module Six is Issues In Grantee Evaluation And Self-Evaluation

- Module Seven is creating a Performance Monitoring and Evaluation Plan

- Module Eight is using statistical analysis at USAID Mali

- Module Nine is about Reporting

MACR●
INTERNATIONAL INC.

TECHNICAL ASSISTANCE
TEAM 1997-1999

Skills Transfer Module 1 Introduction to Monitoring and Evaluation the Basic Concepts, Terms and Tools Used

What is Evaluation?

Evaluation as part of scientific inquiry, has a long history and has been applied to all disciplines and sectors of life (politics, economics, health, education, social services, religion, criminal justice, environmental quality, wildlife management, engineering, agriculture, etc.). A considerable amount of literature is available about evaluation models, how to conduct evaluations and the application and interpretations of various evaluation approaches. Whole disciplines of study have emerged around evaluation and international professional association exists to support the professional practice of evaluation. One could devote their entire career to the study, or application of evaluation processes to real world needs.

USAID defines evaluation as the following:

Evaluation is oriented towards decision making, it is a formal approach to study the degree of achievement towards a desired effect of a specific activity. Evaluation furnishes data to the decision-maker on client (beneficiary) needs, it provides feedback on the activities, strategies and weaknesses of a program.

What 'triggers' Evaluation?

- Monitoring indicates an unexpected result
- Information is needed for key management decisions
- Reviews identify key questions to be resolved
- Customer feedback indicates a problem or concern
- Breakdown in a critical assumption
- Extracting lessons learned

Why Evaluation?

- To understand
  - Why things are progressing as they are?
  - What impact (intended and unintended) is occurring?
  - What actions might be taken either to put things back on track or to revise targets based on experience?
  - For example:
    - Why isn't contraceptive prevalence increasing?
    - Why are girls not attending school?

What is monitoring?

Monitoring is used to determine how well the program is carried out at different levels and at what cost; it tracks changes that occur over time in terms of resource inputs, production and use of services. This differs from evaluation, which is a time-based discipline (the change from time 1 to time 2) or impact assessment that measures the extent to which this change can be attributed to the program intervention (usually at the end of a program or after 5 years). The results of program monitoring are indispensable for program management because they inform the manager whether the program is on track, where the problems are, and what unexpected results have occurred.[3] It is closely related to evaluation in that it can be used to provide feedback that can help correct problems during the course of program implementation.

USAID defines (performance) monitoring as the following:

Performance Monitoring is a process of collecting and analyzing data to measure the performance of a program, process, or activity against expected results. A defined set of indicators is constructed to regularly track the key aspects of performance. Performance reflects effectiveness in converting inputs to outputs, outcomes and impacts.

Why Monitor?

- To predict

- To correct courses of action

- To ensure achievement of results

- To document results

Monitoring and Evaluation

Monitoring and evaluation centers around collecting simple information (qualitative and quantitative) that can be used to develop and implement effective programs, monitor their performance, and ultimately measure their impact. Three types of activities are used to gather this information:[4]

- Program monitoring

- Evaluation research

- Performance or impact evaluation.

It is the responsibility of the Monitoring and Evaluation Specialist and the Results Package Managers to differentiate between which methods are the best to use given the varying programs operating at USAID and the resources available. It is also the responsibility of the M&E Specialists and RPMs to know which methods to use to carry out these operations and what tools exist at USAID (and in the field.) The M&E Plans for each SO Team should assist you in this process. For more information on creating an M&E Plan please refer to *Module 7 Creating a Performance Monitoring and Evaluation Plan*.

---

[3] Bertrand, J.T. R.J. Magnani, and J.C. Knowles. 1994. *Handbook of Indicators for Family Planning Program Evaluation.* Chapel Hill, NC: TheEVALUATION Project.p. 2.
[4] *Child Survival*, BASICS Volume no. 5, Spring 1998 p 2.

Indicators for measuring performance:

An **indicator** is a variable that measures different aspects of a program's data: inputs, process outputs and results. Indicators provide the quantitative and qualitative detail of a set of objectives. They are statements about the situation that will exist when an objective is reached. Suggested selection criteria for choosing indicators are as follows:

- **Measurable**
  - \- # of people who know correct condom use
- **Precise**
  - \- % new employees receiving management training within first 6 months
- **Consistent (reliable)**
  - \- %age of budget allocated to agricultural programs
- **Sensitive**
  - \- % of births assisted by a medically qualified person
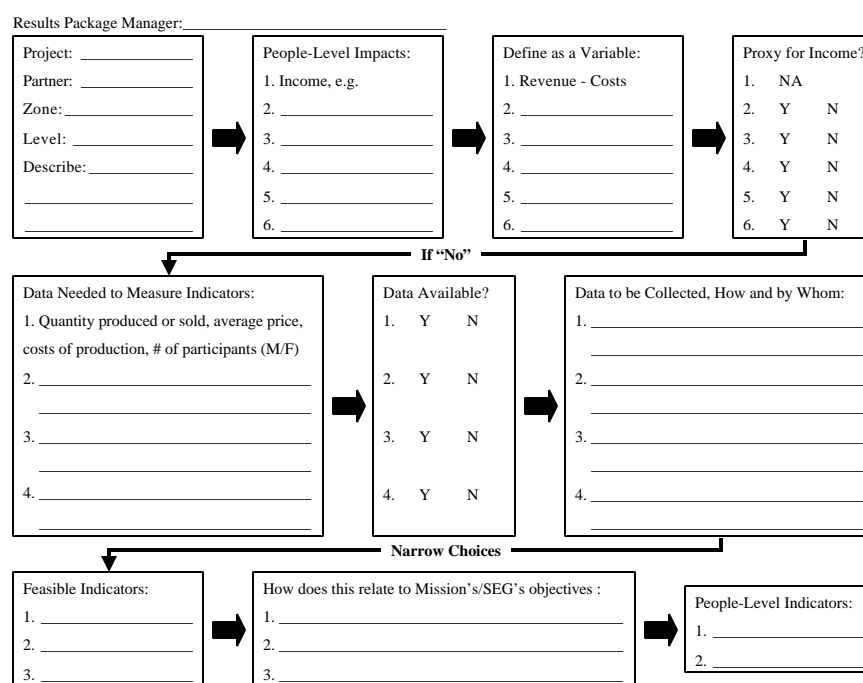
Figure 1.1. Indicator Grid

A grid like this can be used to test whether an indicator meets the above performance measures and if the indicator provides information that leads to overall goals. It is a compliment to the Performance Monitoring Plan because it tests the relevance of a selected indicator and works through the same standards for tracking indicators while adding the concept of data quality and availability. The ADS Performance Monitoring Plan will be revisited in *Module 7 Creating a Performance Monitoring and Evaluation Plan*.

Performance Indicator: A particular characteristic or dimension used to measure intended changes defined by an organizational unit's results framework. Performance indicators are used to observe progress and to measure actual results compared to expected results. Performance indicators serve to answer "how" or "whether" a unit is progressing towards its objective, rather than why/why not such progress is being made. Performance indicators are usually expressed in quantifiable terms, and should be objective and measurable (numeric values, percentages, scores and indices). Quantitative indicators are

preferred in most cases, although in certain circumstances qualitative indicators are appropriate.

Impact: Impact refers to medium or long-term developmental changes. (Some writers also refer to a further class of outcome indicators, more specific to project activities than impact indicators, which may be sectoral statistics, and deal more with the direct effect of project outputs on beneficiaries.) Measures of change often involve complex statistics about economic or social welfare and depend on data that are gathered from beneficiaries. Early indications of impact may be obtained by surveying beneficiaries' perceptions about project services. This type of leading indicator has the twin benefit of consultation with stakeholders and advance warning of problems that might arise.

Process indicator: Measures what happens during implementation. Often, they are tabulated as a set of contracted completions or milestone events taken from an activity plan.

Output indicators: Show the immediate physical and financial output of the project: physical quantities, organizational strengthening, initial flows of services. They included performance measures based on cost or operational ratios.

Results: A broad term used to refer to the effects of a program or project. The terms "outputs", "outcome" and "impact" describe more precisely the different types of results. A results description could include:

- reference number
- results statement
- indicator(s)
- implementers/partners responsible for the result
- timeframe
- possibly, resources

Results also define objectives at several levels:

- National
- Program
- Activities

Therefore, when elaborating indicators define expected results at the:

- National level
- Program level; and
- Activities level

Quantitative versus Qualitative Indicators [5]

More tends to be made of the distinction between qualitative and quantitative data than is warranted. Not everything that is important can be counted, and much that can be counted is not worth knowing. Many of our most meaningful measures of development involve subjective judgments (for example, the capacity of organizations or the transparency of political institutions). Often, these judgments can be objectively operationalized, using surveys or expert panels or well-defined rating systems, and translated into rank-ordered categories, a form of quantifying judgments.

---

[5] CDIE Tips no 12 *Guidelines for Indicator and Data Quality.*

It might be more helpful to think of quantitative and qualitative as a matter of degree. On the quantitative end of the continuum are measures that involve continual, equal-interval scales with true zero points (such as GDP per capita, infant mortality rates, school enrollment rates). In the middle are data for which the frequency of various occurrences can be counted, and categorized, and perhaps even rank-ordered. At the qualitative extreme are data that can be captured only by descriptive narrative. Though USAID collects and uses data from throughout this continuum, most of the performance information collected by operating units probably falls somewhere in the middle. For example, most of the data the Agency collects on policy reform, institutional strengthening, and customer feedback are measured on some type of ranked (ordinal) scale. Such scales, when clearly operationalized, provide a good illustration of how more subjective data can be usefully and effectively quantified. To read more about scales, please refer to *Module 8 Using Statistical Analysis*.

The quantitative-versus-qualitative debate is not an either-or question with some pre-defined best answer. Within the context of USAID's performance-based management systems, the choice of more quantitative or qualitative indicators involves trade-off among practicality and cost, objectivity and comparability, and the directness of validity (meaningfulness) of the measure. Managers need to pick indicators that provide useful information and an adequate level of comparability over time. More meaningful and comparable indicators are the goal, regardless of how "quantitative" or "qualitative" they are.

Notes on Common Indicators

Some people feel that a common indicator is a concept, which if adopted, would be IMPOSED on them from USAID/Washington. Washington can't make you do anything you don't want to do; you decide which indicators you need to use for program management[6].

CDIE has been working on common indicators for the last two years. They have a working group that is looking at the possibility of introducing common indicators at USAID. Last year the common indicators lists were circulated with a memo originated by Melissa Rosser at CDIE. She has promised to update this information when it is available. The working group continues to struggle with the question of common indicators. Common indicators are the answer for future harmonization and streamlining of programs. A selection of common indicators mean fewer headaches for RPMs and program managers in terms of "creating" indicators which 80% of the time is frequently "re-creating the wheel." With common indicators working groups only need to select and come to consensus about which indicators to use.

Why use common indicators?
- Compare data among agencies
- Compare same type of data over years
- Looks at tendencies over years

---

[6] Believe it or not, the decisions to accept what USAID/Washington staff dictate as correct indicators of progress is a conscientious choice made by SO Teams in the field, by giving that decision making power away to someone else.

Table 1.1. What does the indicator measure?

Objective Level	Indicator Type
Goal	Impact
Purpose	Outcomes or effects
Outputs	Outputs
Activities	Process
Inputs	Inputs

An indicator must be

- Measurable
- Precise
- Consistent (reliable)
- Sensitive

For more information on indicators please also refer to:

CDIE Tips no 12 *Guidelines for Indicator and Data Quality.*

CDIE Tips no 6 *Selecting Performance Indicators.*

Chapter: III *Methodological Approach : program Monitoring*. Bertrand, Magnani and Rutenberg. *Evaluating Family Planning Programs.* The Measure Project, September 1996.

Norton, Marley. Chief Evaluation Division. *Guidelines for Data Collection, Monitoring and Evaluation Plans for A.I.D. assisted Projects. A.I.D. Program Design and Evaluation Methodology Report No. 9.* USAID: April 1987

MACR
INTERNATIONAL INC.

Skills Transfer Module 2 Using "Logic Models" To Describe Programs

(Or The R4 is your Friend.)

This skills transfer module addresses these concepts:

- Using "logic models' to diagram programs[7]

- Using a Results Framework in M&E

Generic Chain of Events Logic Model

Program Initiation Factors	Factors Affecting Delivery	Program Activities	Immediate and Inter-mediate Outcomes	Final Program Goals (SO's)

Environments
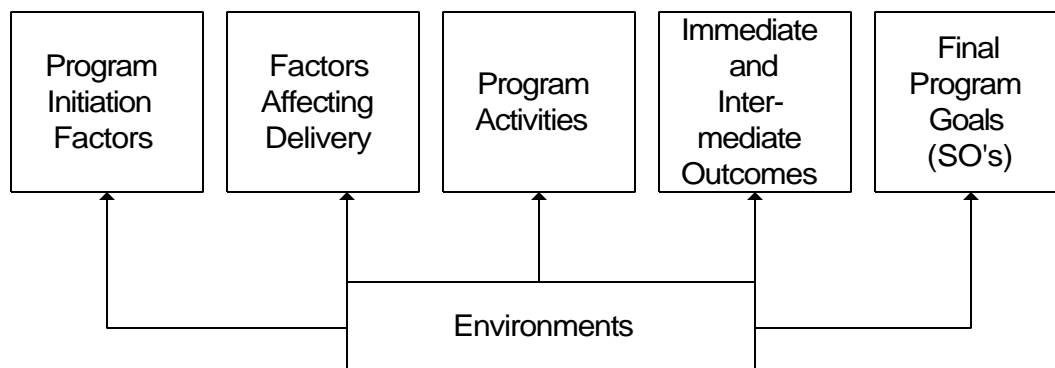
Figure 2.1. The Logic Model as a Program Diagram

Figure 2.1 illustrates a chain of events logic model when diagramming a program. This depicts a natural cycle that most development programs follow, which is why it is called generic. If you will note the presence of intermediate outcomes and final goals is the reason this particular model was chosen to demonstrate the program cycle. Although it is packaged differently, this chain of events model best resembles the USAID paradigm. The USAID paradigm is called a Results Framework and it is a logical framework that pictorially represents how factors of a program feed into the overall USAID objectives. Good Results Frameworks must show logical consistency:

---

[7] Additional reference materials: Hermann, Jack. "Using Logic Models to Strengthen service Program Development" (unpublished paper, Macro International 1996)

- Linkages between Intermediate Results (IRs) and Strategic Objectives (SOs) are <u>causal</u> in nature

- Logical relationship between IRs and SOs is direct and clear where IRs are lower-level results which contribute to SOs

- IRs include key partner as well as USAID-funded results

How the pieces of the Results Frameworks relate to M&E

To identify the parts of a program and the logical connections with assumptions, outcomes, and environments you would use what is called a logical or conceptual framework. At USAID the term Results Framework is commonly accepted as the label for what is otherwise known as a logical framework or a cognitive map. The figure below represents the basic components of a Results Framework. This conceptual framework shows the stages of the program linked to the level of result it is intended to achieve. Results are categorized by project planning stages as: **inputs, processes, outputs** or **outcomes** which are then labeled in the USAID paradigm as Intermediate and Strategic Objective Results.
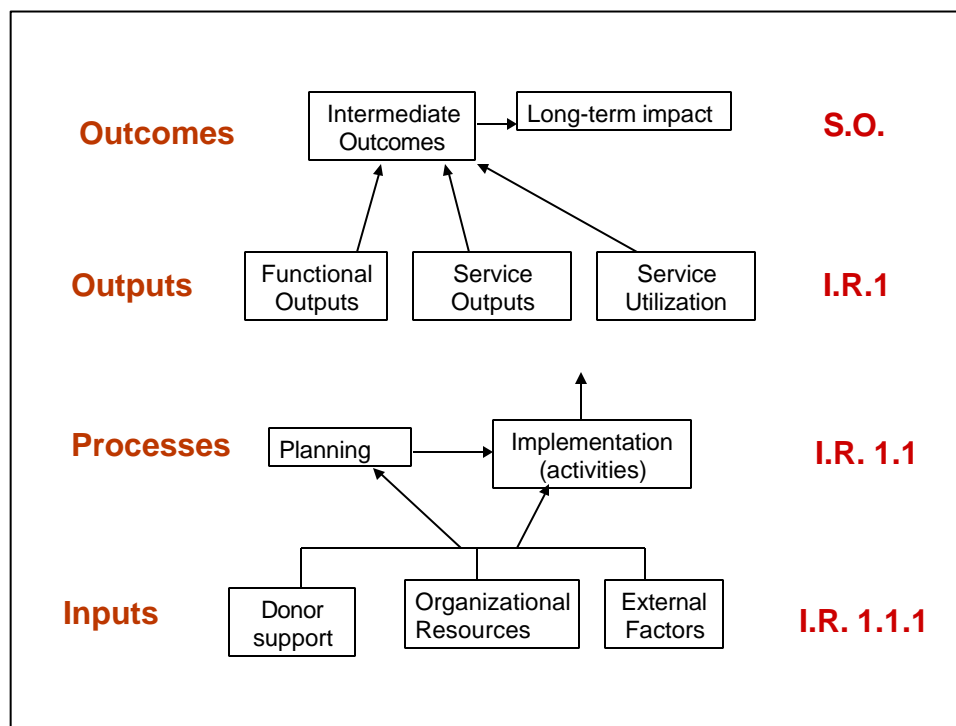
Figure 2.2. Conceptual Framework for a program

Inputs
- What do you need to do your project?
- What types and levels of resources were allocated to this intervention?
- Were qualified personnel available to implement activities?
- What was the unit cost of each resource? The total cost of each resource?
- Data from project records purchase orders, etc.

Processes
- How do you do it?
- Was the workplan designed and agreed upon?
- What activities are being carried out? How many people involved? Frequency? Cost?
- Are reporting mechanisms in place?
- Did data come from project reports?

Outputs
- What does the project do?
- What was the number of persons trained (by category)?
- What was the volume of products sold/disbursed (agricultural products, contraceptives, condoms, etc.)?
- Did the number of service outlets increase (schools, health centers, and radio stations)?
- Did the number of people using the services increase (children in schools, people receiving micro-credit loans)?

Intermediate Outcomes
- What is the goal of the project?
- Was there a change in the key behavior among target population?
- Was there a fundamental change in the political environment regarding the issue?

Long-term outcomes
- Why is the project important to the region/country?

- How will it change lives?

- Was there a social change in the population (lower mortality, fertility)?

- Was there sustainability of the project's goals?

- Was there institutionalization of the project's goals?

Using a Results Framework in M&E

Using a logic model as an evaluation tool helps to:
- Understand the line of reasoning the program uses in assuming that the interventions will have their intended results

- Determine how the project can be evaluated

- Determine what process and outcomes to measure

- Determine if the evaluation correctly assesses the success of the project in achieving stated goals

Steps in the project planning phase...
- Clarify the objectives of the program

- Describe how the program "should" work (the conceptual framework)

- Identify components to be monitored

- Define relevant indicators

- Identify sources of data

Within the conceptual Framework outlined above the various data sources can be conveniently lumped into two categories: Outputs that are program based and Outcomes that are population based. Table 2.1 below delineates the types of data sources for this model

Table 2.1. Identifying Sources of Data

Component to Measure	Source of Data
Outputs Program Based	Program records Service statistics Facility surveys Data on the commercial sector Government sources Special studies
Outcomes Population Based	Census information DHS-type household surveys

Using a logic model as a management tool will:

- Determine the fidelity of program implementation

- Identify the sequence for implementing the program's components

- Monitor program change

- Help explain the program to interested individuals, potential funding sources, new program staff, and the general public

- Makes implementation steps explicit

Helps to assess plausibility of intentions for:

- Theory

- Program delivery

- Measurement

The following example is part of the framework for the SEG SOT's People Level Indicators. It was created as a demonstrative framework showing how the people level indicators would fit into a logical model. This is by no means the Results Framework for the SOT, but these concepts will fall neatly into place in that Results Framework once the team has elaborated the indicators. A results framework is a logic model, but it is not the only type of logic model. Understanding that there are several types of logical frameworks will help when working indicators and Intermediate Results from the PVO's logical frameworks into the larger USAID paradigm.

USAID/SEG Mali S.O. Impact on People-Level Indicators

	SO: Increased Value-Added of Specific Sub-sectors to National Income

National Level

Policy Intervention:	IR1: Increased Production	IR2: Increased Processing	IR3: Increased Trade	IR4: Savings and Credit
Aggregate Effects:	IND 1.1 Volume of Rice Paddy Production IND 1.2 Total Cattle Off-take IND 1.3 Total Area Under Harvest	IND 2.1 Volume of Cereals Processed IND 2.2 Volume of Hides and Skins Processed IND 2.3 Volume of Fruits and Vegetables Processed	IND 3.1 Cereals Exports IND 3.2 Livestock Exports IND 3.3 Alternative Commodity Exports	IND 4.1 Savings Mobilized IND 4.2 Outstanding Loans IND 4.3 Number of Loans
Intermediate Effects:	Income, Prices Food Availability	Income, Prices	Income, Prices	Income, Prices

People Level ~ Quality of Life

Individual Outcomes:	Health, Nutrition Life expectancy & Women's Status

Figure 2.3. The Logic Model for the SEG SO Team's People Level Indicators

In the model above, People Level Indicators are lumped into categories by Intermediate Result. This is different from the Results Framework. This particular logic model was designed to show how the proposed people level indicators would fall under the Strategic Objective and Intermediate Results, this is not a substitute for that Results Framework. This was designed to help the SO Team members understand where the proposed indicators fall into their already existing paradigm. Logic Models come in all shapes and sizes. Their primary purpose is to help us clarify our thought process. Thoughts are messy, so sometimes we need to create a picture to convey the meaning of our words.

PERFORMANCE MONITORING &
EVALUATION PROJECT 1996-1998

Skills Transfer Module 3 A Framework for Planning Project Data Needs

A Framework for Planning Project Data Needs

Project phases and information types

A Framework for Planning Project Data Needs

What types of evaluative information may be useful at various phases of a USAID project?

This framework table lists various types of information and evaluation procedures that should be considered *at each phase of the life of a project* (or in relation to other sets of activities, such as a local sub-project or a complete Results Package. Systematic uses of information include basic statistical information, results, indicators and performance monitoring, as well as occasional, one-time "evaluations." (See additional pages for more information about each information and evaluation type.)

Table 3.1. A Framework for Planning Project Data Needs

Project Phases	Useful Information/Evaluation Types
A. Needs assessment • Focused on circumstances in Mali concerning the target problem.	A.1. Government statistics (baseline?) A.2. Problem analysis using ethnography, focus groups, interviews, etc. A.3. Expert panel A.4. Survey data

Project Phases	Useful Information/Evaluation Types
B. Consultation with partners (in addition to continuing consultation at all phases of a project).	B.1. Examining data from and about partners' activities. B.2. Rapid appraisal methods, such as direct observation, focus groups, and key informant interviews with partners or potential partners.
C. Project design and piloting • What strategies and activities will AID/support in relation to the needs? • SO's and IR's developed or modified	C.1. Consulting prior literature, especially for validated examples of successful strategies. C.2. Site visits to countries with successful strategies (if feasible). C.3. "Formative evaluation" = trial runs, pilot projects, qualitative feedback on proposed strategies, etc. to test potential strategies.
D. Project implementation, by grantees and contractor(s) • Start-up activities • Early delivery/piloting • Full operation(may be in cycles)	D.1. Site-level data collection about activities and ongoing service delivery (e.g., for indicators and internal monitoring). D.2. Assessment of organization development and capacity building factors (by self and /or others) D.3. Monitoring by USAID activity managers D.4. Impact evaluation-(discrete studies) • to follow-up individual participants or trainees • to assess longer term effects • for causal attribution
E. Project close out and termination, with final review and summarization	E.1. Projected analysis of likely sustainability E.2. End of project summary evaluation, by a review team(traditional AID evaluation) E.3. Quantitative data summary and synthesis
F. Sustaining/institutionalizing project components and /or benefits; dissemination of "lessons learned'	F.1. Follow-on analysis of extent of sustainability of project components and benefits (e.g., 6 months after close-out) F.2. Papers, publications and other dissemination activities.

Suggestions for Use of a Variety of Information and Evaluation Types in Relation to Project Phases[8]

A. Phase One-Needs assessment

A period for assessing the circumstances in Mali concerning the target population, what is already being done by government agencies, other donors, and so forth.

Some Useful Information on needs assessment types:

A.1. *Government statistics*-Relevant agencies often have aggregate statistics on national trends, sometimes desegregated by population groups, provinces, etc. Multiple years of data are better than just one, to see what trends were occurring, for example, over the past 5 years. Look also for similar relevant statistics from other countries, for example, in World Bank publications.

Such statistics may later become the baseline for a performance indicator. Again, the baseline is more meaningful if it has more than one data-point, to see if past trends are stable or varied.

A.2*. Rapid appraisal methods*-Problem analysis using ethnography, focus groups, interviews, etc. Emphasis is on understanding how this problem affects the lives of potential participants and beneficiaries and the potential feasibility and effectiveness of possible USAID strategies and activities.

See CDIE TIPS#5 *Using Rapid Appraisal Methods*

A.3. *Expert panel*- Convene a group of experts and stakeholders- such as program managers of similar activities, government officials of relevant agencies, and academics who have done relevant research-to discuss the current status of the problem, relevant alternative approaches to solving it (theory), and the strategic interventions appropriate for USAID.

A.4. *Survey data*- conduct a sample survey (or use existing survey data) to assess the scope and characteristics of the problem in the population and in potentially relevant sectors of it (e.g., women versus men, rural versus urban, various ethnic groups, etc.)

B. Phase Two- Consultation with partners

While consultation should be an on-going process, including "consultation" as a distinct phase emphasizes its importance; this step reinforces the emphases on participatory processes in reengineering guidance. Information sources are taken into account, ensuring greater relevance of the project plan developed, encouraging partner "buy-in" to help make it work, and increases the likelihood that intended results will be achieved.

Some Useful Information/Evaluation Types

B.1. *Examining data from and about partners activities*- partners may have access to relevant information that is not generally available, such as previous internal evaluation reports, access to detailed agency records, and experience with frequently occurring implementation problems.

B.2. *Rapid appraisal methods*, such as direct observation, focus groups, and key informant interviews with partners or potential partners. The focus of the use of rapid appraisal at this

---

[8] Note: All references listed are available in the Documentation Center.

phase would be to find out how potential partners are working on the problem, and their perceptions of needed strategies for solution.

See CDIE TIPS#5 *Using Rapid Appraisal Methods*

C. Phase Three-Project design and piloting

This phase becomes more concrete about what strategies and activities USAID will support in relation to the needs. Information from the previous two phases is used to plan (or re-design) the specific projects and activities that will be included in USAID's budget, including the most appropriate mechanisms to fund them (i.e., contract, grant, or other type of agreement). SOs and IRs would need to be developed or modified as a part of the planning.

Some Useful Information/Evaluation types

C.1. *Consulting prior literature*- it is especially useful to look for previous, validated examples of successful strategies, for the problem you are addressing. Use both Malian and international sources, such as evaluations, etc. Be alert both for successful strategies and unsuccessful ones. For example, it is well documented that providing information alone is not sufficient to change people's behavior.

C.2. *Site visits* to countries with successful strategies (if feasible). For a major new program, it may by desirable for an Activity Manager to visit a location where strategy being seriously considered is being successfully implemented, with positive results that are already documented. This provides a close look at how the details in that strategy are being carried out, and an opportunity to consult in person with project managers who are in daily contact with field-based activities.

C.3. *Formative evaluation* as a final step of project design, it is often desirable to pilot test the intended activities, along with data collection to document their initial results. Assessing the trial run or pilot project is known as "formative evaluation" as it is intended to help "form" the final design of the full project. Often rapid appraisal methods are appropriate here, as they emphasize quick, qualitative feedback on initial strategies.

Rapid Appraisal methods: Quick, low-cost ways to gather data systematically in support of managers' information needs. Examples: focus groups, key informant interviews, community meeting/interviews, direct observations, use of existing data or documents.

For further information, see CDIE Performance Monitoring and Evaluation TIPS from your Monitoring and Evaluation Toolkit, Sept. 98,

- *Number5: Using Rapid Appraisal Methods,*

- *Number 4: Using Direct Observation Techniques*

- *Number 10: Conducting Focus Group Interviews*

- *Number 2: Conducting Key Informant Interviews*

Recommended reading:

Valadez, Joseph and Michael Bamberger, *Monitoring and evaluating Social Programs in Developing countries: A Handbook for Policymakers, Managers, and researchers.* Washington, DC: The World Bank, 1994. Chapter 7- " Simple, Rapid and Cost-Effective Ways to Assess Projects".

D. Phase Four-Project implementation, by grantees and contractor(s):

This phase is the heart of project operations. It often involves the following sub phase: start-up activities, such as assembling all the resources that will be needed, along with securing office space, staffing, staff orientation and training, and sometimes other organizational processes, such as financial management, procedures design, and designing data collection methods for monitoring and indicators. This process may require 3 to 6 months.

- Early delivery/piloting (not done during the prior phase)

- Full operation, when the intended project activities are carried out fully, with intended beneficiaries. Sometime full operation occurs in cycles, such as annual rounds of application for a higher education project.

Some Useful Information/Evaluation Types

D.1. *Site –level data collection about activities and service delivery-*ongoing (e.g., for indicators and internal monitoring).

D.2. *Assessment of organization development and capacity building factors*- If organizational capacity building is a major part of the project, or if implementation factors are complex, a specific assessment of these processes can be done, either by the implementing organization itself and/or by external evaluators. One or more case studies are often used for such assessment (see Skills Transfer *Module 5 Using Case Studies to Illuminate Results*). Multiple methods of data collection are often used for such case studies, such as key informant interviews, document review, observation of program delivery activities, and/focus groups with staff members of clients.

D.3. *Monitoring* by USAID activity managers-on-going information gathering about status of a project. (See *Module4 on Project Monitoring for Activity Managers*.)

D.4. *Impact evaluation-* the term "impact evaluation" is used in many different ways. Some of its meanings are as follows:

- *To follow-up individual participants* or trainees, for example via a systematic survey to assess whether and how they have used the training or other activity funded by USAID.

- *To assess longer-term influences* from an entire project, particularly a project working toward institutional or systemic change. For example, if AID funding contributed to economic policy development, that "might assess whether the policy was later implemented, and even if it led to the intended results for macro-or micro-economic development."

- *For causal attribution*-Sometimes it is important to know whether the activities provided by AID do actually cause the results in the indicators. The descriptive data collected via normal indicators and data collection do not provide firm evidence of cause-and-effect relationships, when they do not include a comparative basis for inferring what would have happened in the absence of the program. In order to infer causal impact, researchers and evaluators have developed *experimental and quasi-experimental designs*. (Much of the evaluation literature uses the term "impact evaluation" as synonymous with methods for establishing causal attribution, but USAID has not emphasized this type of evaluation in the past.)

Experimental Design- a rigorous method to assess the effectiveness (causal impact) of a well-defined "treatment" or set of activities. Requires a priori random assignment of participants to a "treatment" versus a control group, or random assignment among alternative "treatments."

For further information, see:

Wholey, J.s, Gatry H.P., and K.E. Newcomer, *Handbook of Practical Program Evaluation*. San Francisco: Jossey Bass, 1994. Chapter 6- 'Constructing Natural Experiment ' and Chapter 6- "Constructing Natural Experiments" and Chapter 7 - "Convincing Quasi-Experiments: the Interrupted Time Series and Regression-Discontinuity Designs"

Valadez, Joseph and Michael Bamberger. *Monitoring and Evaluating Social Programs in Developing countries: A Handbook for Policymakers, Managers, and Researchers.* Washington, DC: The World Bank, 1994 Chapter 8- "Quasi-Experimental Design for Estimating the Size of Project Impacts."

Orwin, Robert G. "Twenty-one Years and Counting: the Interrupted time Series Comes of Age" in Chelimsky E. and Shadish, WR. *Evaluation for the 21st Century: A Handbook.* Thousand Oaks, CA: Sage, 1997.

Herman, JL., Morris LL, Fitz-Gibbon CT. *Evaluator's Handbook*. Center for the Study of Evaluation University of California, Los Angeles. Sage Publications, 1987.

E. Phase Five - Project close out

Project close out and termination, with final review and summarization. This is the traditional time period for conducting USAID evaluation. It is usually a time period of re-assessing the results of the project, and deciding whether to continue the project activities, and if so, with the same or a different strategy. "Lessons learned" can be extracted to apply to other projects, in the same or other countries. The results of this phase might also be the needs assessment for a next phase of projects.

Some Useful information/ Evaluation Types

E.1. *Projected analysis of likely sustainability* to forecast what parts of the current project are likely to be sustained. This type of information collection might focus on what arrangements are being made for institutionalizing the project activities into on-going organizational responsibilities, rather than being a special, externally - funded project. One might examine the financial resources that will support continued operations, whether continued staffing and staff training will be available, and so forth.

E.2. *End of Project summary evaluation*, by a review team (traditional AID evaluation.) This type of evaluation, usually a synthesis of information by a team of external evaluators may still be extremely valuable. I can document the full project, summarize the data collected during various points in its life cycle, develop summarizing judgements of overall accomplishments, and develop "lessons learned" to be applied to new projects. It usually uses a "mixed method" approach, with multiple data sources, depending on what is already available and relevant to the project.

E.3. *Quantitative data summary and synthesis* - If extensive quantitative data has been collected during the project, for example within the monitoring systems of multiple projects, additional cross-project, synthesis may be useful. This might use statistical methods, for

example, to analyze results in relationship to differences among the multiple sites, in the populations served, in quality or extent of implementation, in types of activities offered, and so forth.

F. Phase Six-Sustaining/institutionalizing project components and/or benefits; and dissemination of "lessons learned"

These activities occur *after* the close - out of the formal project, in order to facilitate the continuation of the project benefits, either for the same or new beneficiaries with similar needs. Further, it is often desirable to disseminate the results of a successful project via reports, publications, conferences, and so forth.

Some Useful Information/Evaluation Types

F.1. *Follow-on analysis of extent of sustainability* of project components and benefits (e.g., 6 months after close-out). This type of evaluative study is similar to the Projection of sustainability in Phase Five, but goes further to examine actual continuation of intended activities. It is frequently done via a sample survey (if large number of potential beneficiaries are involved), or with case study methods (when the project focus has been on institutional change). See *Module 5 Using Case Studies to Illuminate Results.*

F.2. *Papers, publications and other dissemination activities*. Communicating the results of projects- whether successful or unsuccessful-is an important part of learning from others' experience and helping them to learn from this project. This is a key source of evidence for accumulating knowledge about "what works" for various purposes, and to provide the source of prior literature for others to use in the project design phase.

N.B. "Lessons Learned" can also be Recommendations and Best or Better Practices.

MACR◉
I N T E R N A T I O N A L I N C.

PERFORMANCE MONITORING &
EVALUATION PROJECT 1996-1998

Skills Transfer Module 4 Project Monitoring for Activity Managers

Components of this Module:

What is "monitoring"? Intensities of monitoring

Use of indicators data for more frequent monitoring

Site visits-a checklist for assessing organizational capacity

How to use projects' written reports for monitoring

Orientation to Project Monitoring

What is project "monitoring"?

The actions of an Activities Manager to collect information about the on-going status of a project (grant, contractor, or sub-contractor). The purpose is to alert the manager as to whether intended project activities are occurring, and to diagnose the sources of any problems or "slippage" in implementing the intended project schedule.

This is somewhat distinguishable from **performance monitoring** which is regular, on-going information collected to track programs and projects, to alert management as to whether actual results are being achieved as planned. This should include tracking inputs, activity processes, and outputs, as well as results. (From CDIE TIPS #11 *The Role of Evaluation in USAID*)

Project monitoring can occur at various levels of intensity:

- From close, frequent monitoring to infrequent, "hands off" monitoring

- When should an Activity Manager use intensive versus "hand-off" monitoring?

Project monitoring can include diverse types of data collection:

- Regular, frequent examination of indicator data

- Site visits for observation, interviews, inspection of data collection procedures

- Examination and analysis of project documents, such as quarterly or semi-annual reports

Using Indicators Data for Project Management

Development of a Monitoring and Evaluation Plan

- Development of benchmarks for project monitoring

- Data collection methods

- Reporting frequency

- Application of indicators data

- Data reliability and verification

- Responsiveness to monitoring information

- Expected outcomes from implementation of the grant or contract

- Specific activities or processes leading to attainment of expected outcomes

- Benchmarks for indicating progress on implementation

- Data collection plan

  See *Module 7 Creating a Performance Monitoring and Evaluation Plan*

Data Collection Methods

Both kinds of data, quantitative and qualitative, can be collected via:

- Project records

- Direct observations

- Case studies

- Surveys

Frequency of Reporting

Information should be reported back to the activity manager at regularly scheduled intervals. At USAID/Mali the PVOs report every six months and annual statistics are normally on a calendar year. These intervals will be selected by the following criteria:

- Availability of information

- Decision making cycles of the unit

- Type of project

- Delinquency or other problems with delivery

Application of Indicators Data

- Cross section data

- Assessment of the status quo

- Comparative analysis of performance within same project

- Problem identification

- Impact assessment or formative evaluation

Time series data

- Assessment of performance or change over life of project

- Performance against set objectives of project

- Planning-outstanding work on project; other similar projects

- Evaluations

Data reliability and verification

Reliability of data can be confirmed by:

- Checking project records

- Spot checks on site visits

- Interviews with project staff/beneficiaries

- Reliability assessment based on historical information

Responsiveness

A monitoring plan needs to be flexible and able to adapt to the information manager's changing needs. Change in the implementation plan or the monitoring plan is not undesirable, and is often necessary. Activity managers should use monitoring information to respond to lessons learned from the implementation of the project.

Monitoring Checklist

Table 4.1. Template For Assessing Organizational Capacity

Components of Effective organizations	Current Status?
1. Mission/Goals of Organization	
a. Key developers have agreement on and commitment to a well-understood mission and goals for the organization.	
b. The nature and the needs of intended outcomes, specified in measurable terms.	
c. The organization's goals are clearly connected with a set of intended outcomes, specified in measurable terms.	
d. Establishing and maintaining itself as an organization is included as a legitimate part of the organization's mission.	
2. Program Activities	
a. The program activities to be delivered are clearly connected to the overall mission, and to the needs articulated for and by the organization's clients.	

Components of Effective Organizations	Current Status?
b. The effectiveness of the specific program activities has been established by prior research (or the unit includes a research component to rigorously assess the effectiveness of its activities, as delivered.	
c. Actual program activities are feasible in addressing the organization's mission and goals (over- simplification and over- promising are avoided.)	
d. The program activities are comprehensive enough to address key aspects of the needs identified.	
e. Marketing plans and procedures are established to obtain and continue intended client participation.	
f. The scope of program delivery is scaled to the resources and staffing available	
3. Monitoring Plans and Processes	
a. The organization establishes a plan for collecting and using monitoring data that articulate its operational goals	
b. Monitoring data assess whether the clients served are those appropriate for its mission and goals.	
c. Monitoring plans include indicators of success linked to key stages of the program delivery (e.g., in a management information system).	
4. Governance	
a. The organization's legal status (e.g., as a free- standing structure; part of governmental organization; or part of another private organization) is appropriate for its mission and environment.	
b. If a Board of Directors is the governing body for the unit, it focuses on major decisions for broad courses of action, rather than day-to-day operations.	
c. If an Advisory Council is established, it links the unit with other key organization in its environment and/or provides expertise needed by the unit.	
d. Members of a Board of Directors or Advisory Council are chosen for their relevance to that entity, and receive orientation and training as needed.	
e. Procedures for decision-making promote consultation with staff and stakeholders, but also promote timely decisions.	
5. Leadership	
a. A person has overall authority for managing the organization, as well as the charisma to articulate its mission both internally and externally.	

Components of Effective Organizations	Current Status?
b. The key leadership team has skills and experience in the content area of the program components and in both external and internal management	
c. If the organization is new, there is continuity of leadership from the period of establishing the organizational unit through the first several years of operation, although the leadership emphasis may evolve as the organization grows.	
d. The leadership team is able to balance priorities between developing the programmatic mission, creating the necessary administrative mechanisms, and supporting creative growth and contributions among staff members.	
e. If the new unit is part of an existing organization, there is strong support from the leaders of the larger entity.	
f. As the unit grows, leadership is delegated appropriately to sub-division heads for example, in becoming part of the leadership team.	
g. The staff leadership and the Board of Directors remain in harmony, with clearly differentiated roles.	
6. Staffing	
a. The roles needed (both professional and support) to carry out the program components and all together tasks are appropriately identified and related to each other in an organizational	
b. Support staff are included in appropriate relationships to senior staff.	
c. People with needed skills and experience backgrounds are available, recruited and hired in timely manner.	
d. Staff members' skills are effectively utilized and rewarded, both financially and via feedback from supervisors and co-workers.	
e. Effective attention is given to staff diversity and staff development.	
7. Resources/Finances	
a. The financial resources available are appropriate for the stage of organizational development and scope of program activities.	
b. Funding organizations' goals and priorities are fully congruent with the mission of this organization.	
c. Resources are appropriately allocated among program and support components.	
e. Some flexibility in allocation of resources is maintained, with slack for meeting unforeseen contingencies.	

Components of Effective Organizations	Current Status?
f. Space and physical facilities are obtained that ate adequate for the numbers of staff members and the program activities being carried out.	
8. Administrative Processes	
a. Normal operating routines and procedure established, but are not used as rigid rules.	
b. Modern procedures are established for support functions, such as financial management, personnel recruitment and administration, purchasing supplies and equipment, and operating and maintaining equipment, such as computes.	
c. Procedures for trouble-shooting and problem solving are built into the organization's operating culture and expectations.	
9. Communications	
a. Good internal communication channels are established to foster formal and informal communications, both vertically and horizontally.	
b. Administrators actively support staff members who deliver program activities, and utilize contributions of all staff in their areas of expertise.	
c. The sub-divisions of the organization maintain good communications with each other, as the organization grows.	
10. Relationships with Environments	
a. Key leaders maintain strong communications with all-important sectors of environments: clients, government, regulating agencies, and potential partners or competitors.	
b. If part of a larger organization, the unit has appropriate linkages and division of tasks/activities with collaborating entities.	
c. Staff members monitor and contribute to their broader environments, such as the local and national media, and professional organizations.	
d. The unit does not experience excessive disturbance from the environment, such as political controversies or the exit of key supporters.	

MACRO

INTERNATIONAL INC.

PERFORMANCE MONITORING &
EVALUATION PROJECT 1996-1998

Skills Transfer Module 5 Using Case Studies to Illuminate Results

What is a "case study"?

Types and Purposes of case studies

Steps in conduction a case study

What is a "Case Study"?

A case study is a method for learning about a complex instance, based on a comprehensive understanding of that instance, obtained by extensive description and analysis of that instance taken as a whole and in its context.

> From: USGAO/PEMD, Case study
> Evaluations. GAO Transfer Paper 10.1.9 1990.

Some aspects of this definition:

- "about a complex instance" these are often relationships in the case

- "based on a comprehensive understanding" the goal is to obtain as complete a picture as possible of what is (or did) happening in the case

- "obtained by extensive description and influences in the case are indirect or non-linear" multiple factors may affect the outcomes of the case being examined; often focused on "why" or "how" questions

- "obtained by extensive description and analysis" this uses full reach information; this often uses qualitative narrative information; may combine multiple types of data, and compare them to the findings

- "taken as a whole" the focus is on examining one or more instances, such as: an individual (e.g., role of a key political figure), a site of a specific event (e.g., a major demonstration or labor strike), a policy or governmental decision, a specific project or organization, such as sub-grantees delivering a service

- "In its context" this refers to all factors that could affect what happened in that case, such as a decision-making process, political pressures, normal operating procedures of an organization, environmental factors, etc.

Various types of case studies are conducted for diverse purposes. They often complement data collected by other methods, such as use of internal records or surveys as seen in Table 5.1. below.

Table 5.1. Types and Purposes of Case Studies

Type of Case Study	Purpose
A. Illustrative	A. To describe what is happening and why, in one or a few cases; to illuminate what a situation is like; helps make the unfamiliar familiar; example: describing the operations of a typical community-building project.
B. Exploratory	B. To develop the evaluation questions, measures, or analytic strategies for a larger study. To understand a problem area in a needs assessment when considerable uncertainty exists about the background and reasons for the problem.
C. Critical Instance or Case	C. To examine in depth a situation of unique interest, such as a site that is unusually effective or, in contrast, showing results substantially below targeted levels. Or a particular situation may raise allegations of impropriety that need investigation.
D. Assessing Program Implementation	D. To understand the processes underlying program operations and results, especially if implementation problems are suspected. Often conducted in more than one site, with cross-case comparison among sites. It is desirable to obtain information over time, from early in the life of a project.
E. Assessing Project or Policy Results	E. To determine whether the activities undertaken by USAID led to the expected results, particularly in complex situations or institutional change, in which multiple factors influence results. These may complement quantitative indicators.
F. Cumulative Case Studies	F. Brings together the findings from case studies done at different times, to aggregate or compare findings from cases conducted at different points in time. May use quantitative methods to analyze coded data from multiple individual case studies.

A case study should be distinguished from a brief "case example" containing anecdotal descriptive information illustrating a project or result. These are often effective ways of calling attention to a problem or to convey the "flavor" of a result, but are not full case studies.

Steps in Conducting a Case Study

Basic research principles

- Identify the problem, is it our problem or is it the population's problem?

- Define the problem indigenous to the population concerned, and other pertinent factors

- Elaborate the research question from the population and the researcher's perspective

- Chose the methodology to be used whether interview, observation, analysis of text or survey

- Collect and analyze your data

- Write your report

Design

Basic decisions are made about the case study, including:

- What type of case study do we want and for what purpose(s)?

- Who are the audiences for the case study?

- How many cases do we want to include? Of what types?

- What types of data will be collected or obtained from available sources?

- What time frame will be examined? Current status only? Past (history) to present?

- Current and continuing into future, with multiple data collection waves?

- Who will do the case study? With what resources? Contract or grant mechanism?

Site selection (see next page)

If multiple cases will be included, may require assembling preliminary data about the universe of sites, based on such characteristics as size, location, past performance, clients served, funding sources and/or amounts, and so forth, in order to do systematic site selection.

Data collection

Case studies usually use multiple sources of data, (please refer to the list below,) with comparison of multiple sources of evidence to support valid conclusions. Often site visits are used to collect multiple types of evidence.

Types of data and evidence

- Participant Observation, sometimes the observer is also a participant in the events being observed

- Attending meetings or community events, using direct observation

- Conducting structured or open-ended interviews, with a varied range of participants or informed observers of the site/case

- Conducting focus groups, of clients, staff members, community members, etc.

- Examining document archives, such as minutes of meetings, policy memos, lists of persons served, contracts, goals, mission statement, procedures manual (often useful for stating what was intended to happen, the formal system, rather that what does happen)

Questions to ask when collecting data

- Who are the participants?

- What are the social actions that occur in this milieu?

- How does the population think about or speak about these things?

- What social traits must one understand to work with this culture, milieu or setting?

- What types of indigenous knowledge is there that one could uncover?

Obtaining data from computerized systems or other archives

- Collecting physical articles, such as examples of products of new enterprises, physical specimens for health analysis, examples of posters used for public campaign

- Take or collect photographs, for example, of environmental conditions to show physical setting, to show before and after of community development, etc.

Have data collectors pose as clients of the agency being studied, such as a new patient in a clinic, an applicant for a housing loan, etc. Are the services provided efficiently and with the activity components intended?

Table 5.2. How to Select Sites or Instances for Case Studies

Selection basis	When to use and what questions it can answer
Convenience	In this site, selected because it was expedient for data collection purposes, what is happening, and why?
Bracketing	"What is happening at extremes? What explains such differences?"
Best cases	"What accounts for an effective program?"
Worst cases	"Why isn't the program working?"
Cluster	"How do different types of programs compare with each other?"
Representative	"In instances chosen to represent important variations, what is the program like and why?"
Typical	"In a typical site, what is happening and why?"
Special interest	"In this particular circumstance, what is happening and why"
Probability	"What is happening in the program as a whole, and why?"

Data analysis

Most analysis for case studies is qualitative, that is using words, rather than numbers. But quantitative data may be incorporated into the case; for example, when records have been extracted from an information system or when cost data are relevant.

In case studies, planning for data collection, actual data collection, and analysis are often iterative, rather than sequential. In the planning stage, one thinks through the questions and issues that need to be investigated, these then are used and amplified during data collection and analysis often begins during data collection. Sometimes, further data are needed during analysis, as hypothesis or conclusion emerges that requires further evidence.

Reporting

- Usually done as a narrative, organized chronologically or by major issues or themes.

- Most evidence must be presented verbally, rather than as summarizing statistics, so case studies can be lengthy. Good, interesting writing is essential.

Graphic displays are often useful, such as flow charts, critical path charts, tables comparing multiple sites on the same issue or characteristic, and time lines.

The following tables are designed to help you define and trouble shoot your case study. Table 5.3. features the different ways of analyzing case study data to help you keep these concepts separate. Table 5.4. lists common pitfalls in case study evaluation. This is designed to help you look out for potential problems in data collection, design or analysis.

Table 5.3. Ways of Analyzing Case Study Data

Feature	Methodology
Iterative	Data collection and concurrent analysis
OTTR	Observe, think, test and revise
Triangulation	Comparison for multiple, independent sources of evidence before deciding there is a finding
Rival explanations	Developing alternative interpretations of findings and testing through search for confirming and disconfirming evidence until one hypothesis is confirmed and others ruled out
Reproducibility of findings	Establish through analysis of multiple sites and data over time
Plausible and complete	Data analysis ends when a plausible explanation has been developed. Considering completely all the evidence
Specific techniques for handling multi-site data sets	Matrix of categories graphic data displays, tabulating frequency of different events, developing complex tabulations to check for relationships, and ordering information chronologically for time series analysis

Some things to keep in mind when reporting a case study

- The event you are reporting has happened during a specific time frame

- The story should comprehensively tell the perspective of the actors involved, so keep your own perspective separate from those under observation

- Case studies should allow the reader to follow the steps during the transformation of the event under observation

Group interview reports should include the following

- The language, vocabulary and concepts used during the discussion

- The choice of participants and why?

- The principal object of the interview

- What type of notes the interviewer kept

- The analysis method used

- Comparison of different or conflicting view points

- What conclusions can you draw from this experience?

Table 5.4. Some common Pitfalls in Case Study Evaluation

Study stage	Common pitfall
Design	Mismatch between criteria for the specific job and what the case study application can do insufficient attention to contrasts and comparisons needed for purposes of the study
Site selection	More sites selected than needed; fewer sites selected than needed; inappropriate basis for site selection, for the particular job, or evaluation and research questions
Data collection	Reliability jeopardized by lack of common guidance in data collection; lack of quality control in data collector roles and responsibilities: impartiality threatened; overly loose relationship between data collected and the evaluation question: inadequacy of information
Data analysis	Insufficient attention to requirements of analytic plan chosen" low plausibility of results"; insufficient attention to management and data reduction; inefficiency, lateness, incomplete use of data; inadequate methods relating findings across sites; inadequate methods for relating qualitative and quantitative data within sites
Reporting	Over generalization, compared to actual basis for site studied and requirements for inference in the design; inadequate interpretation, un-integrated narrative; results not adequately related to user questions; inadequate attention to threats to impartiality (e.g. Sample bias) and the extent to which these have been avoided

MACR● INTERNATIONAL INC.

Skills Transfer Module 6 Issues in Grantee Evaluation and Self-Evaluation

Discussion of appropriate roles for grantees in various types of evaluation

Advantages and disadvantages of grantee self-evaluation

Some suggestion for increasing the credibility of self-evaluation

CDIE TIPS # 1 *Conducting a Participatory Evaluation*

Appropriate Roles for Activity Managers, External Evaluators and Grantees in Types of Information/Evaluation

This table uses the *Framework for Planning Project Data Needs* from Module 3 (this shows information and evaluation types useful at various project phases). Here the Measure Project added a column with suggested responsibilities for each type of information collection. A major suggestion is that appropriate roles for grantees will differ at different points in time, and for different types of evaluation. "Who should do what?" will also depend on the type of grantee; the skill level of staff members; prior experience with various types of data collection, analysis and use. For example, a grantee that is an academic research center is likely to have different capabilities than a new, small community NGO.

Table 6.1. Suggested Responsibility for Each Project Phase

Project Phases	Useful Information/Evaluation Types	Who Should Be Responsible?
A. Needs assessment -Focused on circumstances in Mali concerning the target problem.	A.1. Government statistics 　2. Problem analysis using ethnography, focus groups interviews, etc. 　3. Expert panel 　4. Survey data	A. Activity Managers, with support from contractors and/or grantees

Project Phases	Useful Information/Evaluation Types	Who Should Be Responsible?
B. Consultation with partners (In addition to continuing consultation at all phases of a project).	B1. Examining data from and about partners activities 2. Rapid appraisal methods, such as direct observation, focus groups, and key informant interviews with partners or potential partners.	B. AM's with active involvement of partners.
C. Project design and piloting What strategies and activities will AID/SA Support in relation to the needs? (SO's and IR' developed or modified)	C.1. Consulting prior literature-especially for validated examples of successful strategies. 2. Site visits to countries with successful strategies if feasible). 3. "Formative evaluation" = trial run, pilot projects, qualitative feedback on proposed strategies, etc. to test potential strategies.	A.M's, and/or grantee. Support from contractors, e.g., for systematic literature review. 3. For example, if grantee is piloting a strategy, it might be responsible for evaluating that strategy.
D. Project implementation, by grantees and contractor(s) Start-up activities Early delivery/piloting Full operation(may be inn cycles)	D.1. Site-level data collection about activities and service delivery-ongoing (e.g., for indicators and internal monitoring). 2. Assessment of organization development and capacity building factors(by self and/or others) 3. Monitoring by USAID activity managers 4. Impact evaluation-(discrete studies) • To follow-up individual participants or trainees • To assess longer term effects for causal attribution	Grantees have major responsibility for data about service delivery/activities, with support from contractors and AM's. If other indicators data needed from other sources, obtained by contractor or AM(?) 2. Org'1 development and capacity building-both grantee and external monitor-AM or contractor 4. Either grantee or external, with support from AM. Depends on availability of technical skills needed, and specific + funding.

Project Phases	Useful Information/Evaluation Types	Who Should Be Responsible?
E. Project close out and termination, with final review and summarization	E.1. (Projected analysis of likely sustainability 2. End of project summary evaluation,(traditional AID evaluation) 3. Quantitative data summary and synthesis	E.1. Probably external evaluator or contractor. External review team (?), especially if analysis is predominantly judgmental. Internal or external, based on skills.
F. Sustaining/institutionalizing project components and/of "lessons learned"	F.1. Follow-on analysis of extent of sustainability of project components and benefits(e.g., 6 months after close-out) 2. Papers, publications and other dissemination activities.	F.1. External contractor or evaluator. 2. Jointly, by AM, Contractor and/or grantees

Table 6.2. Advantages and disadvantage of Summary Assessment Evaluation or Impact Evaluation performed by (NGO) grantees versus external evaluators or contractor:

Factors Contributing to High Quality Evaluation	Grantees	External Evaluator/Contractor
Knowledge of project-level activities and issues	Likely high-an advantage	Lower-Disadvantage; time often needed to learn the project
Technical skills available	May be low-a disadvantage	Higher, or can easily be obtained-advantage
Findings being use for program improvement	Likely high-an advantage	May be harder to get findings to people who made decisions about program- disadvantage
Cost of data collection	May be lower if collecting in own organization-an advantage	May be higher, if done by an external contractor or evaluator
Potential for bias	May be higher, because of grantees' interest in "positive" findings-a disadvantage	Likely lower, incentives are usually to be accurate, rather than to produce either positive or negative findings.
Credibility of findings	May be lower, especially if study methods are not well documented- a disadvantage	Likely higher, because no direct self interest; more knowledge of appropriate methods and procedures.

Issues concerning grantee evaluation and self-evaluation

Audience:

Who should do evaluation may depend on who is the intended audience for the study or evaluation:

- If grantees (or others like them) are the key audience, for example, if evaluation is conducted to improve programs, grantees may be the most appropriate unit to conduct the evaluation. See CDIE TIPS # 1 *Conducting a Participatory Evaluation*.

- If external entities are the major audience, e.g., AID/W or host government, it is usually best to have external evaluator or contractor; especially if evaluation is a summarizing assessment.

Conflict of Interest:

A key issue with grantee self evaluations is that the credibility of the findings may be questioned since the grantee is usually in a "conflict of interest" situation desiring to produce only "positive" findings. This is a well- documented finding in social psychology, to attribute positive happenings to one's own efforts, and negative happenings to *external* influences or forces. Part of "attribution theory". Note that much *scientific study is funded by grant mechanisms*, in U.S. and elsewhere, and findings have high credibility, if properly conducted.

Increasing credibility of grantee findings.

Several strategies are useful for increasing the credibility of evaluation conducted by grantees:

At the start-up of grant or specific study evaluation

- *For performance monitoring data for indicators:* Make sure record keeping or other data collection procedures, computerized information systems (if used), and data aggregation methods are designed appropriately for the data needed from each site. Check implementation periodically to maintain quality control. Grantees should have routine procedures for maintaining quality, such as staff training for all new staff, periodic review of procedures for consistency of record keeping, and sometimes supervisor monitoring or duplicate collection/coding for reliability checks.

- Make sure performing unit has *technical skills* for the type of evaluation to be conducted, e.g., in collecting qualitative data, doing a survey, setting up a system of records, etc. If not, obtain consultation, hire needed staff, or perform the study elsewhere.

- *Use a technical advisory committee-* for example, of external evaluators and knowledgeable content experts- to review and critique the plans for a special evaluation study.

- Use a *competitive mechanism* with technical and peer review of evaluation proposals, if not all grantees will automatically obtain evaluation funding.

At the end of a discrete evaluation study, when the written report is available

- Be sure that the methods are fully documented in reports of findings. The credibility of the *findings* rests on the credibility of the *methods*.

- Have draft reports reviewed by an outside evaluator, before being accepted as final, for review of methods and assurance that are based on the data presented and methods used.

For a completed evaluation project, commission an external evaluators to do a formal *evaluation audit*, in which the methods, original data, analysis procedures, etc. are examined in detail, in comparison with the quality standards for that type of study. (Analogy to a financial audit, in which financial records are inspected in accordance with accounting standards.)

Joint Studies

Consider the possibility of joint studies, with grantee(s) working together with a technically skilled contractor or consultant. Grantees are likely to have major input on issues needing study, to describing the reality of the program, and possibly to data collection, with external unit responsible for overall study design, analysis of data and written report. This strategy can combine many of the advantages of boy the internal and external evaluation listed in table 6.2., and minimize the disadvantages.

For more information on conducting an evaluation for USAID please refer to:

Norton, Marley. Chief Evaluation Division. *Guidelines for Data Collection, Monitoring and Evaluation Plans for A.I.D. assisted Projects. A.I.D. Program Design and Evaluation Methodology Report No. 9.* USAID: April 1987

CDIE's *Results Oriented Assistance: A USAID Source Book* found at the following internet address: http://207.175.11.14/usgov/

MACRO
I N T E R N A T I O N A L   I N C.

TECHNICAL ASSISTANCE
TEAM 1997-1999

Skills Transfer Module 7 Creating a Performance Monitoring and Evaluation Plan

Why Develop a Plan?

- Assessing and adjusting the program's strategy or activities

- Reporting program progress to stakeholders, partners, and customers

- Identifying and sharing successful practices and lessons learned with others

- Planning future strategies

The idea behind an M&E plan is to not limit monitoring of information to the restricted set of indicators identified in a Performance Monitoring Plan. We should encourage and support local partners and national institutions to continue tracking changes in a broader development context, to build up systems that not only meet USAID's requirements but also will serve those organizations' needs over the long term. It serves to promote M&E and related data collection activities as an integral part of program implementation and management, not as a discrete activity driven by the changing needs of USAID.

The M&E Plan considers a variety of data collection instruments, e.g., PRA and informal field surveys, formal survey questionnaires, periodic program assessments, institutional development, formal reviews, case studies, and other special studies. These are all complementary data sources.

It considers the array of contextual information to track biophysical and socio-economic changes and monitor activity-specific impacts. It also encourages PVO and local participation in the collection and analysis of these data. If used regularly it can support local institutions (outside USAID missions) in strengthening their information systems, archiving and managing data, and increasing the accessibility of data to a wide range of partners and stakeholders.

Trying to map out and stick to a performance measurement schedule raises an interesting question. It isn't just measurement, its analytical thinking, and its thinking proactively.

In the Project Planning Phase...

- Clarify the objectives of the program

- Describe how the program 'should' work (the conceptual framework)

- Identify components to be monitored

- Define relevant indicators

• Identify sources of data

In the Evaluation Design Phase:

```
┌─────────────────────────────────────┐
│   Report Findings: Decide what       │
│   evaluation will focus on next      │
└─────────────────────────────────────┘
          ┌─────────────────────────┐
          │   Reach Conclusions     │
          └─────────────────────────┘
        ┌─────────────────────────────┐
        │ Study Facts & Info Collected│
        └─────────────────────────────┘
      ┌───────────────────────────────┐
      │ Use Evaluation Methods Chosen │
      └───────────────────────────────┘
    ┌─────────────────────────────────┐
    │  Collect Materials and Resources│
    └─────────────────────────────────┘
  ┌───────────────────────────────────┐
  │     Draft an Evaluation Plan      │
  └───────────────────────────────────┘
 ┌────────────────────────────────────┐
 │      Form an Evaluation Team       │
 └────────────────────────────────────┘
┌─────────────────────────────────────┐
│        Identify Stakeholders        │
└─────────────────────────────────────┘
┌──────────────────────────────────────┐
│          Decide to Evaluate          │
└──────────────────────────────────────┘
```
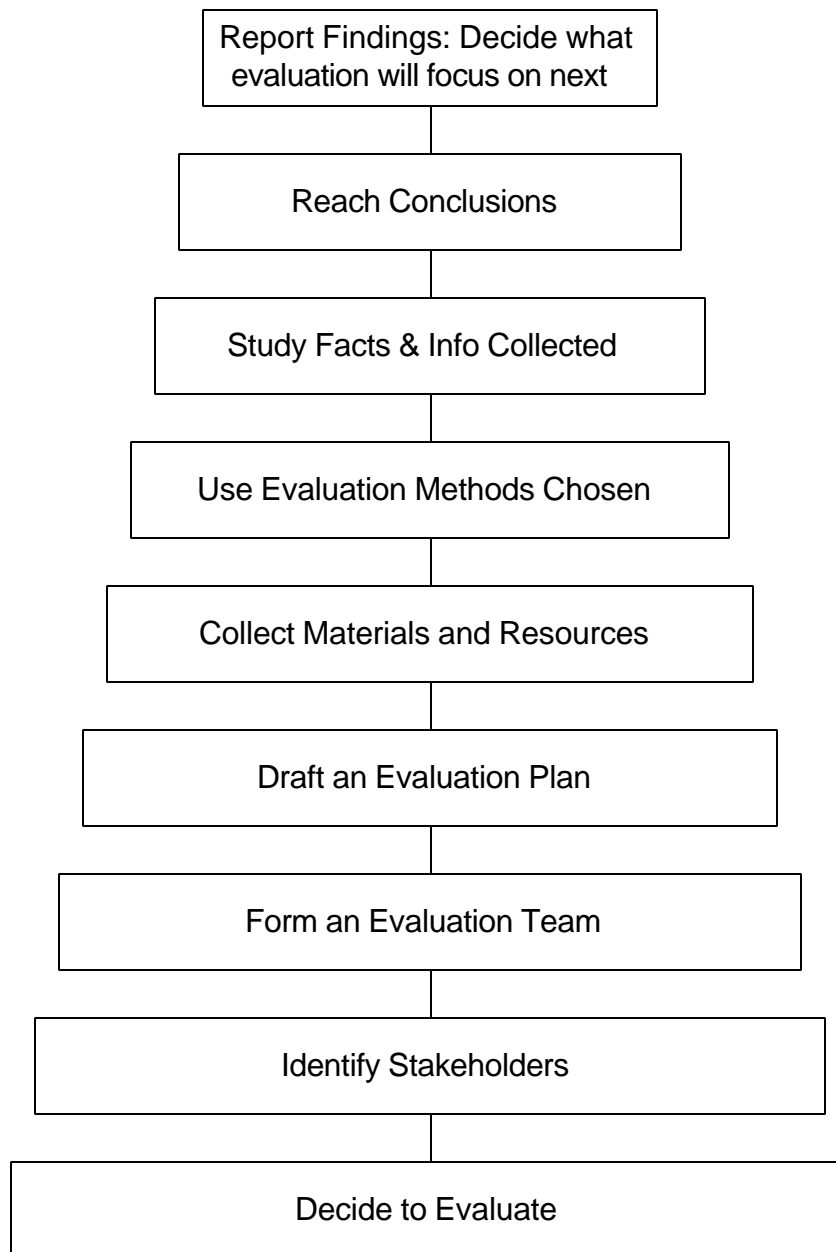
Figure 7.1. A Step Diagram for designing an evaluation

     A project evaluation can be planned in many ways. One good way to get organized and begin the process is to draw a diagram of what you want to do.[9] The step diagram above represents the evaluation process. Starting from the bottom working your way to the top; this diagram illustrates just a few of the steps to be considered when you are planning to conduct

---

[9] Evaluating Your Project. Prepared for the W.K. Kellogg Foundation by the University of Illinois at Chicago, Department of Medical Education March 18, 1994.

an evaluation. This diagram is borrowed from the W.K. Kellogg foundation's manual, *Evaluating Your Project*.[10]

Developing an M&E Implementation Plan

- Define the institutions and individuals responsible for the monitoring and evaluation

- Establish a timetable for specific activities

- Incorporate M&E into the project budget

- Answer the questions in table 7.1. below: The whom? What? Where? Why? and when?

Table 7.1. Summary of an Evaluation Plan

Summary of an Evaluation Plan	
What?	**Scope of an Evaluation** • Goals and objectives of the program • Conceptual framework which defines the links between inputs, processes, outputs and outcomes (results) • Objectives of the evaluation; monitoring/impact assessment
How?	**Methodological Approach** • Study design • Indicators • Sources of data
Who? When? With What Funds?	**Implementation Plan** • Individuals and institutions responsible for different parts of the evaluation • Timetable for specific activities • Budget
Why?	**Dissemination and Utilization of Results** • Target audience • Format and content

Defining the Institutions/Individuals Responsible for the Evaluation

- One institution takes the lead
  - Usually the major stakeholder
- Involve key stakeholders in the planning process
  - All organizations/individuals potentially interested in how the evaluation is carried out, what the results show, and how the information might be used.

---

[10] Prepared for the W.K. Kellogg Foundation by the University of Illinois at Chicago, Department of Medical Education March 18, 1994.

Establishing a timeline for specific activities

COMPONENTS	97/98				98/99				99/00				00/01				01/02			
Develop M&E plan																				
Steps																				
Clarify objectives of program	X																			
Describe how program should work	X	X	X																	
Establish objective of M&E	X	X	X	X																
Identify components to be monitored	X																			
Define relevant indicators	X	X	X	X																
Identify sources of data	X																			
Develop plan for data collection and proce	X																			
Design format for presenting results to Wa	X																			
Design format for presenting results to Partners					X															
Design/Implement Monitoring Activities																				
Routine statistics																				
DNSI National census					X				X				X				X			
MBE data collection	X				X				X				X				X			
MOH data collection					X				X				X				X			
MOH epidemiologic Surveillance	X				X				X				X				X			
PVOs' data collection		X			X		X		X		X		X		X		X		X	
Facility-based surveys																				
Behavioral Sentinal Survey in STI					X															
CScom Survey by AFRICARE			X																	X
Qualitative assessments																				
KAP survey by AFRICARE		X								X										X
Field visits		X				X				X				X						
Special studies																				
Vaccination coverage survey/MOH			X									X								
Youth profiles (Skills Dev + ARH)																				
KPC (population based) survey for CS by SAVE																				
Design Impact Assessment/Evaluation																				
Randomized experiment																				
Quasi-experiment																				
CBD/CS integration survey by Pop. Council					X															
1st cycle test at primary school level					X				X				X				X			
Synergy Eduction/Health experimental study							X	X		X	X	X								
Regression Methods using existing survey																				
DHS survey									X	X	X	X	X	X	X	X				
Utilize Evaluation Results to Modify/improve Programs																				

Figure 7.2. Timeline of Evaluation Activities for USAID/Mali's Youth SO Team 98-02

The figure above represents the Monitoring and Evaluation Plan for the USAID/Mali Youth SO Team for the five years starting in 1998 ending in the year 2002. Note that the Youth Team has conveniently lumped the PVO activities under the types of data collection methods being used. For example the CSCOM survey by AFRICARE is listed under Facility Based Surveys, the type of evaluation activity that categorizes it. More complex scientific studies fall in the last category, notice that the planned DHS is indicated here.

The timeline for specific activities is an M&E plan and should not be confused with the USAID Performance Monitoring Plan listed below. These forms are complementary the "Data Source" and "Scheduled Frequency" categories of the PMP feed into the M&E plan.

Establishing a Timetable for Specific Activities

- For each type of data, specify how often it will be collected and reported.

- Construct a timetable to see the planned progression of the project

- Indicate party(ies) responsible for each type of data

- Define the Institutions/Individuals Responsible for the Evaluation

- Define technical needs and identify sources in-country

- Establish and maintain effective communication channels

The USAID Performance Monitoring Plan (PMP)

An Explanation of a Performance Measurement Tool

Performance monitoring plans are tools used by operating units which serve two principal functions: (1) facilitating the collection and review of *comparable data over time* and (2) supporting the management of the data collection process. The first four elements of a PMP described below - **indicator**, **indicator definition and unit of measurement**, **data source** and **method of data collection or calculation** - are designed to provide managers with enough detailed information to allow for the collection of comparable data from year to year (this information is also very useful when analyzing the relevant data). The last two elements discussed here - **data acquisition** and **analysis, use and reporting** - are meant to furnish managers with the information necessary to manage the internal processes of data collection and use.

Performance Indicator: A performance indicator is a quantitative or qualitative measure used to track progress (or lack thereof) toward achieving a strategic objective or an intermediate result. An indicator is a particular characteristic or dimension of the corresponding result, used to measure intended changes in that result as defined by the results framework and performance targets. A performance indicator should be a precise, direct measure of the corresponding result; it should be practical (i.e., data are available or can be generated when needed at reasonable cost), and it should be disaggregated, where appropriate (by gender, ethnicity, rural/urban, etc.). If the result being measured is focused and appropriately limited, only a few (or even only one) performance indicators are needed per strategic objective or intermediate result. This column should include the indicator statement.

Indicator Definition and Unit of Measurement: These two items are combined into one column, but both aspects are important. State exactly what will be measured. Picture yourself as an evaluation officer who comes in a few years later and needs to know exactly how to replicate collection of the data. What, precisely, is the indicator, and what is the exact unit of measurement?   If the unit of measurement is a percent or a rate, we need to define the numerator and denominator for the indicator. For example, suppose the objective is to increase the practice of contraception. The rough indicator might be the "number of women who practice one or more forms of contraception on a regular basis." How do we define "women" here (age range, only women in union or all women, only women who live in certain geographical areas or in the entire country, etc.).

Table 7.2. Performance Monitoring Plan

Strategic Objective 1 and Related Intermediate Results

PERFORMANCE INDICATOR	INDICATOR DEFINITION AND UNIT OF MEASUREMENT	DATA SOURCE	METHOD/ APPROACH OF DATA COLLECTION OR CALCULATION	DATA ACQUISITION BY MISSION		ANALYSIS, USE & REPORTING	
				SCHEDULE/ FREQUENCY	RESPONSIBLE PERSON(S) AND TEAM	SCHEDULE BY MANAGEMENT EVENT	RESPONSIBLE PERSON(S) AND TEAM
Strategic Objective 1:							
1.	Definition: Unit:						
2.	Definition: Unit:						
3.	Definition: Unit:						
4.	Definition: Unit:						
COMMENTS/NOTES:							

How do we define "forms of contraception?" What do we mean by "on a regular basis?" Are we looking only at the absolute number of women (in which case the unit of measurement would be "number"), or the number as a percent of some whole (which would require definition of both the numerator and denominator). We could use a completely different indicator, e.g., instead of counting women who meet our criteria, we could count person-months of contraceptive use. Other examples: (1) if the indicator is "annual change in grain production," we need to define precisely what we mean by "grain production" (which grains, grown in which regions, etc.) and we need to identify the precise unit of measurement, e.g., metric tons (volume) or US $ (value); (2) for the indicator "number of newly established small enterprises," we need to define both the principal measure and the threshold we will use to define small enterprise (e.g., a small enterprise is an enterprise with fewer than 25 employees or a small enterprise is an enterprise with annual revenues of less than US$ 5,000). We also need to define the criteria we will use to determine whether an enterprise is "established" (e.g., the enterprise is in operation, or is in operation for some minimum period of time, or the enterprise is "registered").

Data Source: Exactly where will the mission get the data? From whom and through what mechanism (a report, a survey, etc.)? Will the data simply be extracted from the monthly reports of extension agents to a project office? Will the data come from a specific question on an annual survey of households, or from a quarterly report from the Ministry of Finance? Again, be as specific as possible. For instance, if the report has a number, give it. If a specific table in a report is the data source, provide this information also. Simply identifying an institution (e.g., the Bureau of Statistics) does not provide an adequate level of detail to ensure comparability.

Method/Approach of Data Collection or Calculation: Think about replication when filling out this column. Will the information in this column permit a newcomer a few years from now to collect data that could be consistently compared with data collected earlier? Are there any details that should be noted? If so, do so. Note that this column also refers to "data calculation." Operating units frequently derive or calculate indicator values by using formulas. Describe the formulas here. In addition to facilitating the collection of comparable data, information in this column is also very useful when interpreting and analyzing the data that have been collected.

While "Data Source" (the previous column) should provide the specifics of the source (e.g., Table 10.4 of the Ministry of Planning and Development's quadrennial report of its Rural Household Budget Survey), "Method/Approach" should provide details on both the data collection process and relevant characteristics of the data. For example, "the Rural Household Budget Survey is a national survey of a random sample of heads of households in all rural communities with populations of less than 500" (though more information than this would be required to fully understand the data). If you need more space for description, use a footnote and write in the Comments/Notes box at the bottom. Operating units that have sponsored or conducted their own primary data collection activity may need to pay particular attention to this column. In some cases it may not be feasible to include all pertinent methodological information (i.e., sampling methodology, survey design, etc.) in the space available in the table. In such situations, reference should be made in the PMP to where the additional methodological information can be found (e.g., in a file with technical notes from the relevant surveys).

Data Acquisition by Mission: "Acquisition" refers to the actual arrival of the data in the Mission. Depending on the data source, this can mean one of two things: Mission staff themselves are responsible for collecting data at their source, or the Mission is receiving data

collected by someone outside the Mission (partners, government counterparts, NGOs, contractors, etc.).   In either case, this column indicates who at the Mission is responsible for ensuring that data are actually available at the Mission, and how often (e.g., annually) and when (e.g., in December) those data are supposed to come into the possession of Mission staff.

Analysis, Use and Reporting: Once data have been collected, they are available for analysis and, in turn, to be used to both inform managers' decisions and to report on performance. The next to last column on this table indicates when various Mission reports are due, or when other management events requiring performance data will occur (e.g., operating units' internal performance reviews). The last column is used to indicate who is responsible for producing or managing these reports or events. Taken as a whole, the information in these columns allows managers to monitor operating units' internal processes for analyzing and using performance information.

Comments/Notes: Use as you wish. This may be the place to document key assumptions being made in the choice of specific indicators and means of data collection, so that the next person will be able to understand. CDIE TIPS #7 *Preparing a Performance Monitoring Plan* revised: 8/8/97[11]

Budgeting the Evaluation

- Identify resources for evaluation in the lead organization that are already covered by other sources and will be available at no extra cost

- Estimate costs for:

    - personnel and other direct costs

    - individual data collection activities

    - data processing and analysis

    - dissemination of results

For more information on Designing an Evaluation Plan, Establishing a Timeline for Specific Activities or Preparing a Performance Monitoring Plan see also:

CDIE Tips #7 Preparing a Performance Monitoring Plan, 1996

Bertrand, Magnani and Rutenberg. *Evaluating Family Planning Programs.* The Measure Project, September 1996. Chapter V "Developing an Implementation Plan."

Bertrand, Magnani and Rutenberg. *Evaluating Family Planning Programs.* The Measure Project, September 1996. Chapter VIII "Summary: Checklist of Steps in Designing an Evaluation Plan."

Lesson Learned

Budgeting for M&E Lessons Learned from USAID Senegal

It is quite difficult to come up with a formulaic number for the expected costs of M&E, i.e., whether it should be 9 or 10 percent, or 5, or whatever. There's a difference between wanting to build M&E capacity for the long term and simply walking in with pre-defined

---

[11] The Role of USAID Partners in Performance Measurement: Challenges and Responses in Collecting Performance Data Notes from a CDIE Brown Bag Lunch Discussion May 6, 1999

performance indicators and wanting to buy data to support those indicators on a regular basis. So, it all depends on what you expect from your partners.[12]

It's hard to draw a clear line between simply getting performance monitoring data for USAID needs and doing general development activities that we want to support along with other donors and partners. Is it just a performance monitoring issue, or should we also be providing support for monitoring long-term environmental changes, or for doing demographic health surveys, etc.? We need to consider both levels. For example, in the case of the $80 million project in Senegal, they had the resources to hire one full-time person dedicated to M&E. That was reasonable. But at the same time, the mission has been committed over the years to work with a regional institution to build its capacity to do M&E.[13]

If you can separate out the costs of doing performance measurement, you haven't got the idea yet. Your performance measurement has to be so deeply built into all the steps of the process, *including* the initial stages of strategic planning, and certainly coming into whatever contracting awards you are making, that it is simply the way that business is done. Evaluation then becomes something that is used on an as-needed basis to explore issues that come up and need to be addressed (e.g., are we achieving things, should we be going faster, and so on). With performance measurement and meeting new data requirements, one of the real problems we have is that we have an existing substrate of contracts that need to be adapted. But, certainly, all contracts that are awarded from some point out need to have built into them thoroughly the acquisition and use of data from the very bottom to the top. As the data move up to higher levels of analysis and reporting, the amount gets smaller and smaller, but the smaller amount of data should be a subset of the data collected at the lower level.[14]

Data Collection Lessons Learned from USAID Senegal

As we struggle to collect appropriate data, we also have to struggle with the question of how those data are going to be organized, archived, analyzed, and managed. What kinds of information systems do we need to establish, and how do we build the capacity to maintain those systems on an ongoing basis?[15]

- Do not limit collection of baseline information to the restricted set of indicators identified in a Performance Monitoring Plan. We should encourage and support local partners and national institutions to continue tracking changes in a broader development context, to build up systems that not only meet USAID's requirements but also will serve those organizations' needs over the long term.

- Avoid losing sight of longer-term development challenges and incremental progress in achieving them, including changes in policy and institutional capacity. We should avoid being overly focused on achieving and reporting on short-term results, which are more easily measurable.

---

[12] Jim Fremming, Senior Program Associate, Management Systems International: The Role of USAID Partners in Performance Measurement: Challenges and Responses in Collecting Performance Data Notes from a CDIE Brown Bag Lunch Discussion May 6, 1999

[13] Bob Winterbottom, Senior Manager, Environment and Natural Resources, International Resources Group The Role of USAID Partners in Performance Measurement: Challenges and Responses in Collecting Performance Data Notes from a CDIE Brown Bag Lunch Discussion May 6, 1999

[14] Nancy McKay, Chief, Division of Project and Food Policy, AFR/DP, and former USAID Representative to Guinea-Bissau The Role of USAID Partners in Performance Measurement: Challenges and Responses in Collecting Performance Data Notes from a CDIE Brown Bag Lunch Discussion May 6, 1999

[15] Bob Winterbottom,

- Use a variety of data collection instruments, e.g., PRA and informal field surveys, formal survey questionnaires, periodic program assessments, institutional development index, legislative reviews, Geographic Information Systems and satellite imagery, case studies, and other special studies. These are all complementary data sources.

- Collect an array of contextual information to track biophysical and socio-economic changes and monitor activity-specific impacts. Also encourage local participation in the collection and analysis of these data.

- Promote M&E and related data collection activities as an integral part of program implementation and management, not as a discrete activity driven by the changing needs of USAID.

- Encourage local ownership in the strategic planning and decision-making processes that are organized to review and analyze performance data. Useful data, e.g., using a Performance Monitoring and Evaluation Plan (above), help stimulate this involvement as people review past experience and look to future challenges.

- Support local institutions (outside USAID missions) in strengthening their information systems, archiving and managing data, and increasing the accessibility of data to a wide range of partners and stakeholders.

- Our implementing partners are getting quite sophisticated about performance indicators and data collection. We have an opportunity here, when issuing an RFP, to get back from them information regarding what *they* consider appropriate indicators for getting the kind of data we need. We should ask for that information as part of our criteria for evaluating proposals, so we can take advantage of what's going on out there when we are establishing a results-based contract.

MACR

INTERNATIONAL INC.

TECHNICAL ASSISTANCE
TEAM 1997-1999

Skills Transfer Module 8 Using statistical analysis

Part 1: Uses and Definitions of Statistics

Part 2: Summary of Statistical Concepts

Part 3: Frequencies using SPSS or any other Statistical Package

Part 4: Descriptives using SPSS or any other Statistical Package

Part 5: Secondary Data[16]

Part 1: Uses and Definitions of Statistics[17]

* **Statistics** are useful for Identifying relationships between sets of empirical data.

* **Statistics** are used for theoretical modeling of the processes that generate data.

* **Statistics** serve to explain aspects of our existence.

   Key questions to ask before launching into statistical analysis:

* What objects are being measured?

* On which variables?

* On what scale is the variable being represented?

   Descriptive v Inferential

* **Descriptive** are direct exhaustive measurements of population characteristics or direct measurement of population parameters.

* **Inferential** are concerned with measuring the characteristics of only a sample from the population and then making inferences or estimates about the corresponding value of the characteristic in the population from which the sample was drawn. These measurements of sample statistics are in turn treated as estimates of population parameters.

---

[16] The Role of USAID Partners in Performance Measurement: Challenges and Responses in Collecting Performance Data Notes from a CDIE Brown Bag Lunch Discussion May 6, 1999
[17] These definitions and the explainations for these concepts were derived from the following source: Kachigan, Sam. Statistical Analysis: An Interdisciplinary Introduction to Univariate and Multivariate Methods. Pp. 3-59.
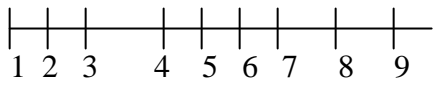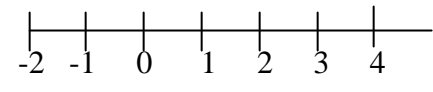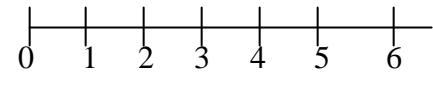
- **Data Reduction** is reducing large sets of data to a smaller summary of the data. In data reduction detail and nuance are sacrificed for parsimony and salience.

- **Error:** All measures are subject to error. **Inferential** measures are subject to the **degrees of confidence** we have in the accuracy of the data.

　　Causal v Correlational

- **Associations** are **descriptive** or **correlational**. No conclusions about **causality** can safely be made about associations

- **Causal** relationships are **experimental** relationships; experimentally based so that researchers can manipulate the levels of one variable to observe changes in another.

- **Descriptive** or **correlational** relationships are observed as they occur in a natural environment.

- In the study of **multivariate analysis** one must be able to distinguish between correlational and experimental relationships.

　　Classes of variables: Variables are objects such as units or people, variables are abstract objects until we have identified the scale by which the objects are measured. Below are four basic types of measurements of scale.

Table 8.1. Classes of Variables

Labels	Description
☻　✡　⌘　☺ 1　2　3　4　5	**Nominal Scale:** Numbers act as labels only, indicating differences in kind (such as identification numbers)
1 2 3　4 5 6 7　8　9	**Ordinal scale:** Numbers represent rank ordering. Differences between ranks have no meaning (such as student standing)
-2 -1　0　1　2　3　4	**Interval scale:** Equal differences between values represent equal amounts, but ratios have no meaning because of the arbitrary location of the zero point (such as temperature)
0　1　2　3　4　5　6	**Ratio scale:** Equal differences between values represent equal amounts. Equal ratios of values are also equivalent because of a genuine zero point (such as a weighted scale)

- **Nominal** scale variables are **categorical** or **qualitative.**

- **Ordinal** scale can be **qualitative** or **quantitative**.

- **Interval** scales are considered **quantitative**.

- **Ratio** scales have **transitive properties** of ordinal and interval scales and are considered **quantitative**.

Qualitative v Quantitative

- **Qualitative** are scaled variable measures on the non-metric nominal and sometimes ordinal scales.

- **Quantitative** are those measured on the metric interval or ratio scales (for example Income is quantitative and religion is qualitative)

Values

- **Discrete** takes on a finite number of values such as the number wrong on a math test, and is usually an integral 1, 2, 3…

- **Continuous** can potentially take on any value such as height, weight, age, grade point average.

- **Dichotomous** values are true and false, male or female.

Different statistical operations take on different variable labels it is important to be able to distinguish between these types.

Part 2: Summary of Statistical Concepts

How can you summarize the various responses people give to a question? [18]

- A frequency table tells you how many people (cases) selected each of the responses to a question. For each code, it contains the number and percentage of the people who gave each response, as well as the number of people for whom responses are not available.

- If you find codes in the frequency table that wasn't used in your coding scheme, you know that an error in data coding or data entry has occurred.

- A count can be transformed into a percentage by dividing it by the total number of responses and multiplying by 100.

- A cumulative percentage is the percentage of cases with values less than or equal to a particular value.

- Pie charts and bar charts are graphical display of counts.

- A histogram is a graphical display of counts for ranges of data values.

- The mode is the data value that occurs most frequently.

- The median is the middle value when data values are arranged from smallest to largest.

- Percentiles are values below which and above which a certain percentage of case values fall.

---

[18] The defnitions of these concepts and ideas for these exercises were borrowed from the following source: Norusis, Marija J. SPSS 6.1 Guide to Data Analysis. Pp 43-129

How can you study the relationship between two or more variables that have a small number of possible values?[19]

- A crosstabulation shows the numbers of cases that have particular combinations of values for two or more variables.

- The number of cases in each cell of a crosstabulation can be expressed as the percentage of all cases in that row (the row percentage) or the percentage of all cases in that column (the column percentage).

- A variable that is thought to influence the values of another variable is called and independent variable.

- The variable that is influenced is called the dependent variable.

- If there is an independent variable, percentages should be calculated so that they sum to 100% for each category of the independent variable.

- When you have more than two variables, you can make separate crosstabulations for each of the combinations of values of the other variables.

- Bar charts can be used to display a crosstabulation graphically.

Table 8.2. Example of a Crosstab

Job or Housework * Total Family Income in quartiles Crosstabulation

			Total Family Income in quartiles				
			24,999 or less	25,000 to 39,999	40,000 to 59,999	60,000 or more	Total
Job or Housework	Very satisfied	Count	53	90	74	110	327
		% of Total	7.1%	12.0%	9.9%	14.7%	43.8%
	Mod satisfied	Count	93	79	61	87	320
		% of Total	12.4%	10.6%	8.2%	11.6%	42.8%
	A little dissatisfied	Count	24	17	14	19	74
		% of Total	3.2%	2.3%	1.9%	2.5%	9.9%
	Very dissatisfied	Count	4	8	7	7	26
		% of Total	.5%	1.1%	.9%	.9%	3.5%
Total		Count	174	194	156	223	747
		% of Total	23.3%	26.0%	20.9%	29.9%	100.0%

What can you say about a population, based on the results observed in a random sample?[20]

- When you take a sample from a population, you won't get the same results as you would if you had data for the entire population.

- The sampling distribution of a statistic tells you, for a particular sample size, about the distribution of all possible sample values of that statistic.

---

[19] The defnitions of these concepts and ideas for these exercises were borrowed from the following source: Norusis, Marija J. SPSS 6.1 Guide to Data Analysis. Pp 43-129

[20] The defnitions of these concepts and ideas for these exercises were borrowed from the following source: Norusis, Marija J. SPSS 6.1 Guide to Data Analysis. Pp 43-129

- From the sampling distribution of a statistic, you can tell if observed sample results are unusual under particular circumstances.

- As the sample size increases, the variability of statistics calculated from the sample decreases.

- The observed significance level is the probability of observing a sample difference at least as the large as the one observed when there is no difference in the population.

- A binomial test is used to test the hypothesis that variable comes from a binomial population with a specified probability of an event occurring. The variable can have only two values.

Before discussing these concepts further, the variables being used in the example above and all subsequent examples below need to be explained. The list of variables used for these processes are listed below in Table 8.3: Variable names and descriptions were selected from the 1993 General Social Survey. This was a survey conducted in 1993, based on a random sample of respondents in the United States. Thus the name, General Social Survey. The data from this survey was distributed by SPSS and is available on their web page for use demonstrating statistical concepts.

Table 8.3. Variables from the 1993 General Social Survey

Variable Name	Description
Age	Age of respondent in years
Sex	1=Male, 2=Female
Educ	Years of education
Income 91	Total family income in 1993 (classified into one of 21 income categories)
Wrkstat	Work status (1=full-time work, 2=Part-time work, 3=Temporarily not working, 4=Unemployed (laid off), 5=retired, 6=In school, 7=Keeping house, 8=Other)
Richwork	"Would you continue or stop working if you became rich?" (1=continue, 2=Stop)
Satjob	Job satisfaction (1=Very satisfied, 2=Moderately satisfied, 3=A little dissatisfied, 4=Very dissatisfied)
Life	"Do you find life exciting, pretty routine, or dull?" (1=Dull, 2=Routine, 3=Exciting)
impjob	"How important to your life is having a fulfilling job?" (1=One of the most important, 2=Very important, 3=Somewhat important, 4=Not too important, 5=Not at all important)

Part 3: Frequencies using SPSS or any other Statistical Package[21]

The Frequencies procedure provides statistics and graphical displays that are useful for describing many types of variables. For a first look at your data, the Frequencies procedure is a good place to start. For a frequency report and bar chart, you can arrange the distinct values in ascending or descending order, or order the categories by their frequencies. In SPSS the frequencies report can label charts with frequencies (the default) or percentages.

Example. "What is the importance to R of having a fulfilling job?" From the output, you might learn that out of the 1500 people in the General Social Survey, 316 people gave the response one *of the most important* that is 21.1% of the 1500 people. Similarly the 238 people who gave the response *somewhat important* are 15.9% of your sample. The 7 people who answered *don't know* are 0.5% of the total sample. (The actual percentage is .47% but SPSS rounds percentages to one decimal place.)

Statistics and plots. Frequency counts, percentages, cumulative percentages, mean, median, mode, sum, standard deviation, variance, range, minimum and maximum values, standard error of the mean, skewness and kurtosis (both with standard errors), quartiles, user-specified percentiles, bar charts, pie charts, and histograms.

Data. Use numeric codes or short strings to code categorical variables (nominal or ordinal level measurements).

Assumptions. The tabulations and percentages provide a useful description for data from any distribution, especially for variables with ordered or unordered categories. Most of the optional summary statistics, such as the mean and standard deviation, are based on normal theory and are appropriate for quantitative variables with symmetric distributions. Robust statistics, such as the median, quartiles, and percentiles, are appropriate for quantitative variables that may or may not met the assumption of normality**.**

Using SPSS to Create a Sample Frequencies Output

➢ From the menu choose:

Statistics

Summarize

Frequencies...

➢ Optionally, you can:
- Click Statistics for descriptive statistics for quantitative variables.
- Click Charts for bar charts, pie charts, and histograms.
- Click Format for the order in which results are displayed.

In the following example both the frequency table and the bar charts options were selected.

To Obtain Frequencies and Statistics
➢ Select one or more categorical or quantitative variables.
➢ In our example we selected the variable "How Important is a fulfilling job?"

---

[21] The defnitions of these concepts and ideas for these exercises were borrowed from the following source: Norusis, Marija J. SPSS 6.1 Guide to Data Analysis. Pp 43-129

Table 8.4. Frequency Table of the Importance of Having a Fulfilling Job?

Importance to R of Having a Fulfilling Job

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	One of most important	316	21.1	21.4	21.4
	Very important	833	55.5	56.3	77.7
	Somewhat important	238	15.9	16.1	93.8
	Not too important	62	4.1	4.2	98.0
	Not at all important	30	2.0	2.0	100.0
	Total	1479	98.6	100.0	
Missing	DK	7	.5		
	NA	14	.9		
	Total	21	1.4		
Total		1500	100.0		

Importance to R of Having a Fulfilling Job
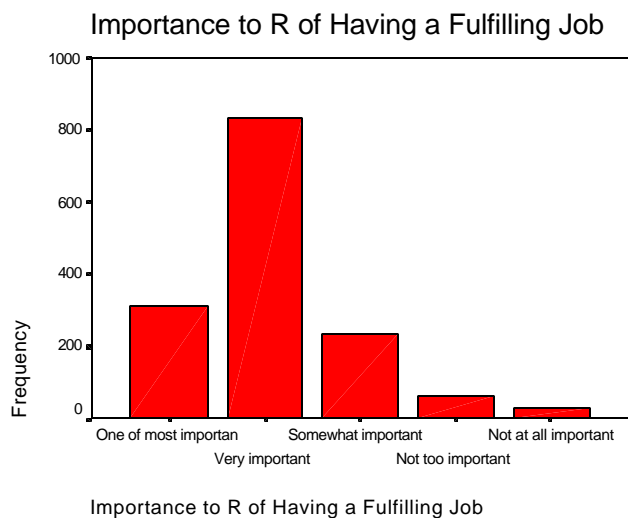
Figure 8.1. Bar Chart of Importance of Having a Fulfilling Job?

Both the frequency table and the bar chart from the example above tell the same story.

Frequencies Statistics[22]

Percentile Values. Values of a quantitative variable that divides the ordered data into groups so that a certain percentage is above and another percentage is below. Quartiles (25th, 50th, and 75th percentiles) divide the observations into four groups of equal size. If you want an equal number of groups other than four, select Cut points for "n" equal groups. You can also specify individual percentiles (for example, the 95th percentile, the value below which 95% of the observation fall).

---

[22] The defnitions of these concepts and ideas for these exercises were borrowed from the following source: Norusis, Marija J. SPSS 6.1 Guide to Data Analysis. Pp 43-129

Table 8.5. Percentile Values for the Age of the People who Responded

Statistics

Age of Respondent

N	Valid	1495
	Missing	5
Percentiles	25	32.00
	50	43.00
	75	59.00

- **Central Tendency**. Statistics that describe the location of the distribution include the mean, median, mode, and sum of all the values.

- **Dispersion.** Statistics that measure the amount of variation or spread in the data include the standard deviation, variance, range, minimum, maximum, and standard error of the mean.

- **Distribution.** Skewness and kurtosis are statistics that describe the shape and symmetry of the distribution. These statistics are displayed with their standard errors.

Table 8.6. Statistical Values for the Age of the People who Responded

Statistics

Age of Respondent

N	Valid	1495
	Missing	5
Mean		46.23
Std. Error of Mean		.45
Median		43.00
Mode		28[a]
Std. Deviation		17.42
Variance		303.39
Skewness		.500
Std. Error of Skewness		.063
Range		71
Minimum		18
Maximum		89

a. Multiple modes exist. The smallest value is shown

Values are group midpoints. If the values in your data are midpoints of groups (for example, ages of all people in their thirties are coded as 35), select this option to estimate the median and percentiles for the original, ungrouped data.

Frequencies Charts[23]

Chart Types: A pie chart displays the contribution of parts to a whole. Each slice of a pie chart corresponds to group defined by a single grouping variable. A bar chart displays the

---

[23] The defnitions of these concepts and ideas for these exercises were borrowed from the following source: Norusis, Marija J. SPSS 6.1 Guide to Data Analysis. Pp 43-129

count for each distinct value or category as a separate bar, allowing you to compare categories visually. A histogram also has bars, but they are plotted along an equal interval scale. The height of each bar is the count of values of a quantitative variable falling within the interval. A histogram shows the shape, center, and spread of the distribution, A normal curve superimposed on a histogram helps you judge whether the data are normally distributed. Below you will find a histogram representing the age of the respondents from the General Social Survey, out of the 1495 people who responded to this question 181 people are between 37.5 and 42.5 years of age.

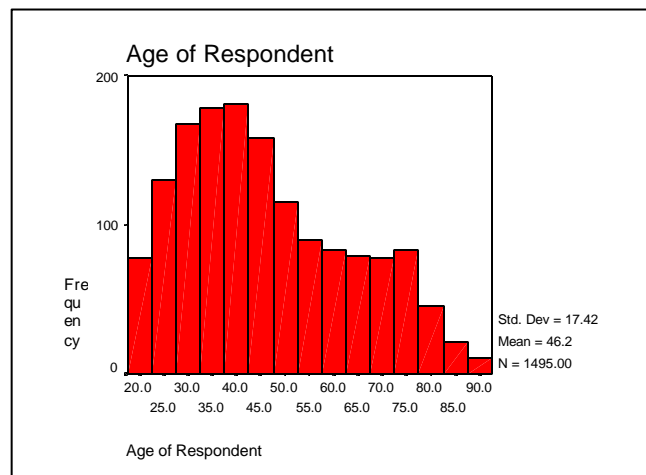**Figure 8.2: Histogram of The Age of Respondents**

Age of Respondent

Chart Values: For bar charts. The scale axis can be labeled by frequency counts or percentages.

Frequency Format

Order by: The frequency table can be arranged according to the actual values in the data or according to the count (frequency of occurrence) of those values, and in either ascending or descending order. However, if you request a histogram or percentiles, Frequencies assumes that the variable is quantitative and displays its values in ascending order.

Multiple Variables. If you produce statistics tables for multiple variables, you can either display all variables in single table (Compare variables) or display a separate statistics table for each variable (Organize output by variables).

Suppress tables with more than "n" categories. This option prevents the display of tables with more than the specified number of values.

Part 4: Descriptives using SPSS or any other Statistical Package[24]

The Descriptives statistics procedure displays univariate summary statistics for several variables in single table and calculates standardized values (z scores). Variables can be ordered by the size of their means (in ascending or descending order,) alphabetically or by the order in which you select the variables (the default).

When z scores are saved, they are added to the data in the Data Editor and available for charts, data listings, and analyses. When variables are recorded in different units (for

---

[24] The defnitions of these concepts and ideas for these exercises were borrowed from the following source: Norusis, Marija J. SPSS 6.1 Guide to Data Analysis. Pp 43-129

example, gross domestic product per capita and percentage literate), a z-score transformation places variables on a common scale for easier visual comparison.

> **Example:** If each case in your data contains the daily sales totals for each member of the sales staff (for example, one entry for Bob, one for Kim, one for Brian, etc.) collected each day for several months, the Descriptives procedure can compute the average daily sales for each staff member and order the results from highest average sales to lowest.

> **Statistics**. Sample size, mean, minimum, maximum, standard error of the mean, and kurtosis and skewness with their standard errors.

> **Data**. Use numeric variables after you have screened them graphically for recording errors, outliers, and distributional anomalies. The Descriptives procedure is very efficient for large files (thousands of cases.)

> **Assumptions.** Most of the available statistics (including z scores) are based on normal theory and are appropriate for quantitative variables (interval-or ratio-level measurements) with symmetric distributions (avoid variables with unordered categories or skewed distributions). The distribution of z scores has the same shape as that of the original data therefore, calculating z scores is not a remedy for problem data.

To Obtain Descriptive Statistics

➢ From the menus choose:

Statistics

> *Summarize*

> > *Descriptives…*

➢ Select one or more variables.

Table 8.7. Example of Descriptive Statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age of Respondent	744	19	82	40.41	11.16
Valid N (listwise)	744				

➢ Optionally, you can:

- Click save standardized values as variables to save z scores as new variables

- Click Options for optional statistics and display order.

Mean and Sum. The mean or arithmetic average is displayed by default.

Dispersion. Statistics that measure the spread or variation in the data include the standard deviation, variance, range, minimum, maximum, and standard error of the mean.

Distribution. Kutosis and skewness are statistics that characterize the shape and symmetry of the distribution. These are displayed with their standard errors.

Display Order. By default, the variables are displayed in the order in which you selected them. Optionally, you can display variables alphabetically, by ascending means, or by descending means.

Part 5: Secondary Data[25]

Table 8.8. Common Problems/Issues with Using Secondary Data

Issue	Problem
Data validity and reliability:	The data do not reflect reality on the ground.
Data Configuration and precision:	The data are not in a form that is useful to the manager/evaluator.
Timeliness	The data are not available at intervals appropriate to the manager's/evaluator's needs.
Access -- short and long-term:	The manager/evaluator is not able to get and use the data throughout the duration of the program.

What is secondary data?

Secondary data is from a source, normally a report, where the data are not collected directly by the person reviewing the information. Secondary data can come in the form of reports, such as PVO report data, or annual national statistics generated by the government line ministries. Whenever the person analyzing the data is not the primary researcher, out in the field collecting the data, the data they are reviewing in report form is considered secondary. Primary data is the actual "data set" that the primary researcher collects. That person can input that data into a software program or transcribe the notes they have taken (qualitative). They then analyze that data to prove or disprove their research assumptions, hypothesis and theories. The data set used for the examples in parts 1-4, above came from a subset of a primary data set. The information you read in the examples (above) has already been analyzed and interpreted, therefore your knowledge of the matter is second hand, therefore secondary.

20 Questions to Ask When Assessing The Usefulness of Secondary Data[26]

General questions that raise red flags and provide context

1) If similar data are available from other source, are they consistent with the data under review, i.e., external consistency?

2) Are the data internally consistent, i.e., when summed, do subtotals equal totals, or, are there any large unexplained variations in the data from one period to the next? (Numerical error raises questions of overall validity.)

3) For what purpose and to answer what questions were the data originally collected?

---

[25] The Role of USAID Partners in Performance Measurement: Challenges and Responses in Collecting Performance Data Notes from a CDIE Brown Bag Lunch Discussion May 6, 1999
[26] The Role of USAID Partners in Performance Measurement: Challenges and Responses in Collecting Performance Data Notes from a CDIE Brown Bag Lunch Discussion May 6, 1999

Data Collection and Analysis

4) What method was used to originally collect the data (e.g., formal survey, observation, remote sensing, informal survey, interviews, self-reporting, etc.)?

   NOTE: If data were collected by some method other than a formal survey, it is still important and appropriate to consider the representativeness of the data.

For Formal Surveys - When Probability Sampling is Used

5) Did every unit (individual, household, firm) in the target population have an equal chance of being selected?

   Related to question #5

6) Is the sampling frame (i.e., the list of units in the target population) up to date?

7) Is the sampling frame comprehensive (and for area frames, are the geographic segments mutually exclusive)?

8) Is the procedure for drawing the sample truly random, including replacement (e.g., simple random, cluster, sequential - with non-ordered sampling frame, etc.)?

For Formal Surveys - When Probability Sampling is Not Used?

9) For Data collected through self-reporting instruments (e.g., mail-in surveys) -- what proportion of the targeted units actually provided information?

For Any Survey?

10) Were the enumerators well trained? How were they trained? Was there any candidate de-selection or other quality control? Were the enumerators insiders or outsiders?

11) Was care taken to minimize the effect of the potential for personal bias the enumerators may bring to the exercise? (Were any of the survey questions "cooked" or leading to a certain type of response?)

12) Did incentives exist for respondents to provide incomplete or untruthful information, whether it be for economic/financial reason (taxes), social/cultural reasons, mistrust of the enumerator or because the respondent was trying to please the enumerator?

13) Were the questions in the survey/questionnaire clear, direct and easy to understand? (If you don't get to see the questionnaire to verify the questions you can't be sure of the quality of the responses.)

14) For self-reporting instrument, were adequate instructions provided to the respondents? (This is a source of considerable survey error.)

15) Were all units in the intended sample contacted and asked for information? If not, was there a systematic or non-random exclusion of units? (With some reliable system the data will not be representative.)

16) Were the raw data transferred, transcribed, organized and analyzed in a careful and appropriate manner? (Each time data is handled the chance for error increases.)

17) Are the data currently in a form/format which will meet the needs of the manager or evaluator? (If not, is it possible to reconfigure the data or get access to the raw data? (With access to the raw data, the analyst can possibly cross-reference data categories in order to test for validity and deepen the analysis.)

Timeliness and Access

18) Does USAID have, or can it get, access to the data? Is it reasonable to expect continued access for the duration of the program?

19) How often are the data collected? Does this meet the needs of the manager or evaluator? (Is data collection consistent -- data collected differently can't be compared easily.)

20) Is there any reason to believe the data will not continue to be collected in accordance with the planned schedule, e.g., the past track record of anticipated institutional or budgetary changes?

MACR◉
INTERNATIONAL INC.

TECHNICAL ASSISTANCE
TEAM 1997-1999

Skills Transfer Module 9 Reporting

Introduction to Reporting

Description and Analysis - writing about Quantitative Data

Description and Analysis - writing about Qualitative Data

Interpretation and Judgement

Note: This skills transfer module is intended as a report writing supplement to *The R4 Guidance Package* included as an **appendix** to this guide.

Reporting[27]

After designing an evaluation and collecting data, the information must be described, analyzed, interpreted, and a judgment made about the meaning of the findings in the context of the project. This process can be complicated and, at times, technical. In fact, many books are dedicated to each method that can be employed in an evaluation. Thus, it is not possible for an introductory manual, such as this, to summarize and adequately explain the necessary skills and techniques of analysis and interpretation. In the following pages, however, some simple techniques and ideas of how to start organizing the data are given. A few ideas are also given to help you begin thinking about how and to whom you will report the results of the evaluation.

> Evaluation data are never completely clear-cut and absolute; studies are always flawed in some way, and there are always questions of reliability and validity. Error-free instruments do not and cannot exist in the measurement of complex human, social, behavioral and psychological phenomena. Then of what good is all this?
>
> Evaluation is only of use if one believes that some systematic information is better than none.
>
> Evaluation does not provide final answers, but it can provide direction.
>
> **Michael Quinn Patton**

The steps of analyses and reporting are often the best time to make extensive use of a consultant or advisor. While you may not use the consultant or advisor to a great extent until this time, they should be at least minimally involved throughout the process to assure continuity and guarantee that the data once collected, can be

---

[27] Evaluating Your Project. Prepared for the W.K. Kellogg Foundation by the University of Illinois at Chicago, Department of Medical Education March 18, 1994.

analyzed in a productive manner.

Description and Analysis

Quantitative Data

With quantitative data (i.e. numbers) you can begin by converting findings (from utilization records, or answers on questionnaires) into percentages or averages. A percentage is a part of something in relation the whole. Calculate a percentage by dividing the number of people or things in a group by the total number in that group and multiply by 100. Percentages are used to compare one number with another. For example, one project may try to implement a new program for teenagers; two ideas have been put forth and it is the job of the two activities would be better accepted and utilized by teenagers in community. The evaluator went to high schools and teenage hangouts in the community and interviewed 128 (randomly sampled) teenagers. She found that 93 of the teenagers preferred plan A, rather than plan B. When presenting this information to the project director the evaluator could use a percentage.

Total number of people interviewed = 128

Total number of teenagers preferring plan"A"=93

$$= \frac{93}{128} * 100$$

=76% prefer plan "A"

An average of a group of numbers is the sum of all numbers in the group divided by the total number of all numbers in the group, then divided by the total number of cases (in the group). Averages are used to compare different groups of numbers using just one number for each group. For example, an evaluator might want to know if certain clients (i.e., preteens) have more or less visits to the project for medical care than other clients (i.e., older teens). The evaluator searches through project records and counts the number of times each preteen and older teen has come to the center for medical treatment in the last month. After going through the entire visit log she totals together all the numbers of visits for each preteen.

Let's assume that only ten pre-teen clients attended the center during the month; the number times they visited were as follows:

5, 8, 7, 4, 9, 3, 6, 8, 9, 11.

Number of preteens=10

Sum of their center visits=59

Therefore, the average of the preteen center visits = 5.9 visits

---

If the same procedure were followed for the older teens and it was found that in a month the older teens average number of visits was 3.2, the difference between the two averages might be useful information for both program planning and recruitment. This is just the beginning of what can be done with groups of numbers. There are many more sophisticated

---

analyses that use statistical procedures. Be aware that statistics is a complex and specialized field and you may wish to use a consultant for advice with the analysis of numerical data.

Qualitative Data

Information gathered from interviews and observation is often recorded in lengthy narratives. Some of these accounts are useful and can stand alone as case studies which illustrate particular points you may want to make in an evaluation report. However in most instances you will want to look for similarities across several accounts, interviews, and /or documents. Usually this involves categorizing your notes into topics that are of interest. This is often done through cutting up a transcript copy of the interview or observation notes and putting pieces on index cards (always remembering to identify where they came from) or by using computer programs that perform the same task electronically. After this organizational task is complete it is easier to locate patterns develop new hypotheses or test hypotheses derived from other sources. As in the section previously discussed, there are many subtle nuances in the analysis of qualitative data that cannot be discussed in this manual. For those interested in using qualitative analysis to strengthen their evaluation findings there are many resources in the documentation center and on line.

Interpretation and Judgement

Interpretation involves looking beyond the mounds of raw data to ask important question about what the results mean, why certain findings resulted, and whether the findings are significant. Making judgements about the merit or worth of results introduces values into the interpretation. Results may be judged positive or negative, good or bad, and point out desirable or undesirable aspects of a program.

Answers for these important large questions should be developed with the collaboration of an evaluation team, and perhaps with other stakeholder groups as well. Involving stakeholders in this part of the evaluation will decrease anxiety. You will gain insight from the stakeholders' knowledge about the program and initiate excitement for evaluation work.

While analyzing and interpreting both quantitative and qualitative data, be careful to avoid the following pitfalls:

1. *Assuming that changes are caused only by the existence of the project.* Several factors, some of which are unrelated to project activities, may be responsible for health status changes in a community. Does the methodology allow the evaluator to discover the specific impact of the project apart from all other factors in the community?

2. *Forgetting that the same evaluation method may give different results when used by different people, or that respondents may tell the evaluator what they believe he/she wants to hear.* For example, two interviewers may ask the same questions but because one was more patient or friendly, the answers they received were different. Real problems or difficulties may be passed over or quickly hidden because people want the project to succeed or appear to be succeeding.

3. *Making a bad choice of groups to compare.* This includes comparing groups that are different in too many ways. For example, it has become evident that research on health and disease among men may not be applicable to women. Therefore, you should not assume that findings hold true for women without taking gender into account.

4. *Claiming that the results of a small-scale evaluation also apply to a wide group of people or to a wide geographic area.* For example, it is misleading to evaluate clients' responses

to midwives in a particular urban area and then claim that the results apply to the country as a whole.

Reporting the Findings

Evaluation findings can be used to document what is happening or has happened in a project, to improve project functioning, account for funds that were spent, convince diverse audiences of the importance of the project, and generate further support for the programs. Evaluation findings presented in the media (e.g., newspaper, radio, etc.) can disseminate information on project activities and the conditions of the target group, thereby increasing local understanding and involvement.

Be creative and innovative by using a variety of techniques including, but not limited to: visual displays, oral presentations, summary statements, interim reports, and informal conversations. Multiple strategies are available for reporting evaluation findings:

- Write and disseminate a complete evaluation report, including an executive summary and appropriate technical appendices.

- Write separate executive summaries targeted at specific audiences or stakeholder groups.

- Write a carefully worded press released put out to the media by a prestigious office or public figure.

- Hold a press conference in conjunction with the press release.

- Make verbal presentations to select groups. Include demonstration exercises that actively involve participants in analysis and interpretations.

- Construct professionally designed graphics, charts, and displays for use in reporting sessions.

- Stage a debate or advocacy-adversary analysis of the findings in which opposing points of view can be fully aired.

Tips for Writing Reports

Almost every evaluation will involve the preparation of some form of written report. When writing reports, there are several tips to keep in mind:

1. *Know who your audience is and what information they need.* Different audiences will want different information, even if they are addressing the same issue.

2. *Relate evaluation information to decisions.* Reports written for decision-making purposes should first state the recommendation and then summarize the relevant evaluation findings.

3. *Start with the most important information.* While writing, imagine that your audience will not have time to read the whole report; you need to be brief, yet informative. You can develop concise reports by writing a clear abstract and starting each chapter, subsection, or paragraph with the most important point.

4. *Highlight all import points.* Significant information can be emphasized with boxes, different type styles, underlining, or CAPITAL LETTERS.

5. *Make your report readable.* Do not use professional jargon or vocabulary that may be difficult to understand. In addition, you should use active verbs as much as possible. Generally, active verbs shorten sentences and increase their impact.

6. *Edit your report*. When you edit your report, look for unnecessary words and phrases. In addition, you should try to write short paragraphs, each covering only a single idea. Often, it is better to have someone else edit your work; if you must edit yourself, allow a day or two to pass between writing and editing.

Examples:

Passive verb sentence:

The rate of pregnancy among adolescents in the program was lower than that of the adolescents in the control group. (20 words)

Active verb sentence:

Adolescents in the program had a lower rate of pregnancy than the control group. (14 words)

In summary, there is no single best way to present evaluation data. The needs and characteristics of the people who will use the information should determine the content and form of the evaluation.

This section is borrowed from: *Evaluating Your Project, Analyzing Your Data,* Prepared for the WK Kellogg Foundation by the University of Illinois at Chicago, Department of Medical Education.

For more information on writing reports please refer to:

The R4 Guidance Package, included as an appendix to this document,

The Revisions to USAID's Performance Planning, Monitoring and Reporting System, also included as an appendix to this document,

Herman, JL., Morris LL, Fitz-Gibbon CT. *Evaluator's Handbook*. Center for the Study of Evaluation University of California, Los Angeles. Sage Publications, 1987.

APPENDICES