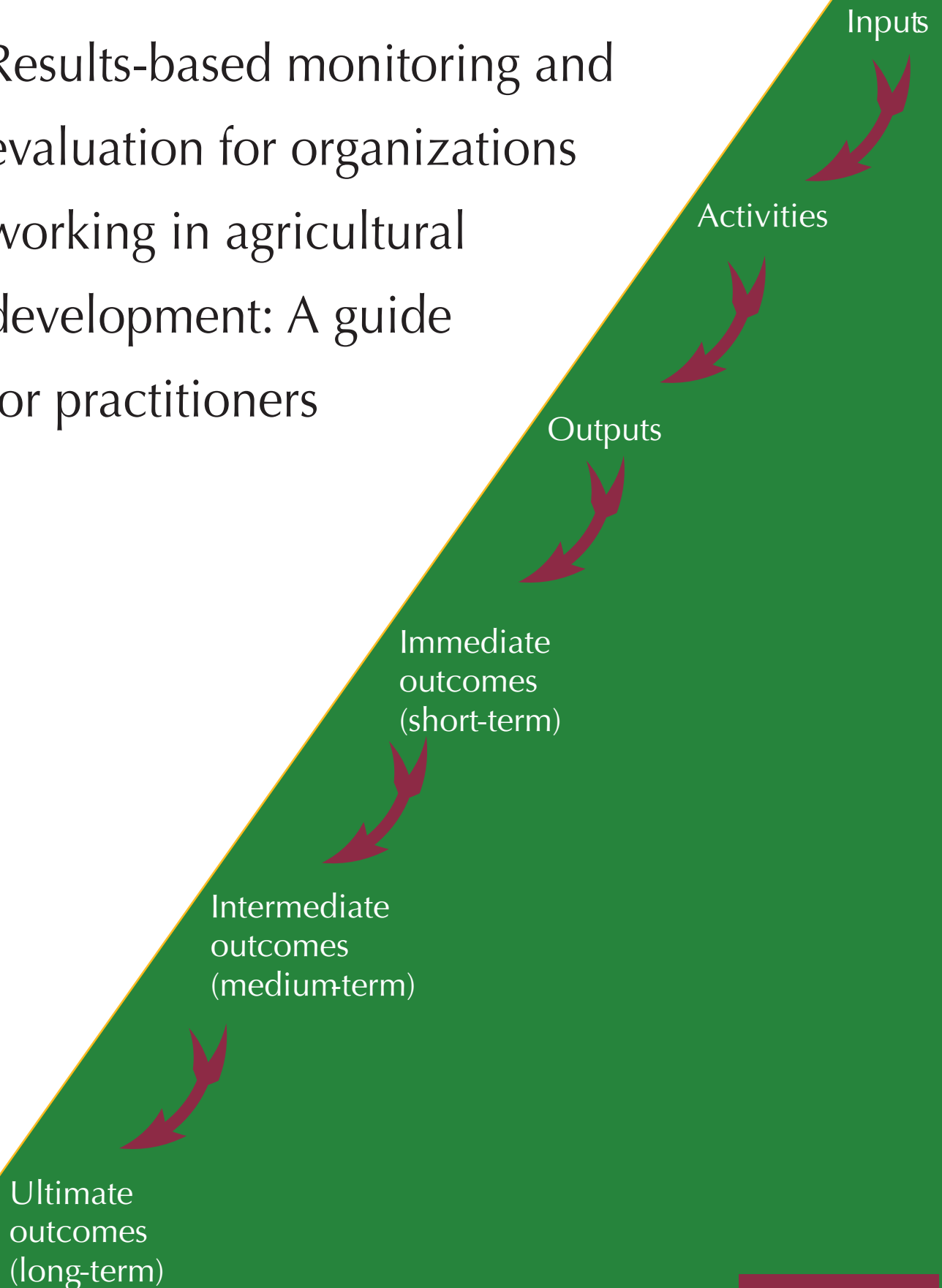


# Results-based monitoring and evaluation for organizations working in agricultural development: A guide for practitioners



# Results-based monitoring and evaluation for organizations working in agricultural development: A guide for practitioners

*Berhanu Gebremedhin, Abraham Getachew and Rebeka Amha*

Improving Productivity and Market Success (IPMS) of Ethiopian Farmers Project,  
International Livestock Research Institute (ILRI), Addis Ababa, Ethiopia

**ILRI**  
INTERNATIONAL  
LIVESTOCK RESEARCH  
INSTITUTE



## Authors' affiliations

Berhanu Gebremedhin, Scientist, Agricultural Economist, IPMS (Improving Productivity and Market Success) of Ethiopian Farmers Project, ILRI

Abraham Getachew, Monitoring and Evaluation Officer, IPMS, ILRI

Rebeka Amha, Monitoring and Evaluation Expert, IPMS, ILRI

© 2010 ILRI (International Livestock Research Institute)

All rights reserved. Parts of this publication may be reproduced for non-commercial use provided that such reproduction shall be subject to acknowledgment of ILRI as holder of copyright.

Editing, design and layout—ILRI Publication Unit, Addis Ababa, Ethiopia.

ISBN 92-9146-245-4

Correct citation: Berhanu Gebremedhin, Abraham Getachew and Rebeka Amha. 2010. *Results-based monitoring and evaluation for organizations working in agricultural development: A guide for practitioners*. ILRI (International Livestock Research Institute), Nairobi, Kenya. 75 pp.

## Table of Contents

List of Figures	v
Abbreviations	vi
Acknowledgements	vii
1 Introduction	1
2 Results-based monitoring and evaluation	3
2.1 Results-based management	3
2.2 Results chain or logical model	3
2.3 Performance framework	6
2.4 Performance measurement framework	8
2.5 Implementation-focused vs. results-based M&E	9
2.6 Uses of M&E information	10
3 Relationships between the concepts and practices of monitoring and evaluation	13
3.1 Monitoring information to support result-based management (RBM)	13
3.2 Evaluative information to support results-based management systems	15
3.3 Differences and complementarities of monitoring and evaluation	21
4 Participatory monitoring and evaluation	23
4.1 Nature of participatory monitoring and evaluation	23
4.2 Uses of PM&E	24
4.3 Fundamental principles of PM&E	25
4.4 Implementing PM&E	26
5 Selecting results (outcomes) to monitor and evaluate	28
5.1 Qualities of good result statements	28
5.2 Process of selecting result statements	28
6 Selecting key performance indicators	32
6.1 Characteristics of good performance indicators	33
6.2 Types of indicators	34
6.3 Constructing indicators	36
7 Setting baselines and targets	38
7.1 Setting baselines	38
7.2 Setting targets	38
8 Data collection and analysis	40
8.1 Source of data and frequency of data collection	40
8.2 Data collection methods and instruments	41
8.3 Choosing data collection method	45
8.4 Preparations for data collection	45
8.5 Data analysis	46

9	Reporting and using monitoring and evaluation information	49
9.1	Reporting M&E information	49
9.2	Using M&E information	50
10	Institutionalizing and sustaining the RBM&E system	52
10.1	Institutionalizing RBM&E	52
10.2	Sustaining RBM&E	52
10.3	The importance of incentives and disincentives in sustaining RBM&E systems	53
10.4	Challenges in institutionalizing and sustaining RBM&E systems	54
	References	56
	Annex 1 Glossary of common M&E terms (OECD 2002)	57

## List of Figures

Figure 1	Results chain of CIDA 1996 (a) and of CIDA 2008 (b)	4
Figure 2	Results chain	5
Figure 3	Building a performance framework	8
Figure 4	Performance measurement framework	9
Figure 5	Integrating results-based M&E with implementation-focused M&E	11
Figure 6	Links between the major objectives of M&E	12
Figure 7	Outcome statements derived from identified problems or issues, IPMS project	31
Figure 8	Methods used for data collection	42

## Abbreviations

ANOVA	Analysis of Variance
BoARD	Bureau of Agriculture and Rural Development
CBO	Community Based Organization
CIDA	Canadian International Development Agency
DAC	Development Assistance Cooperation
HDI	Human Development Index
HIV/AIDS	Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome
IDRC	International Development Research Center
IFAD	International Fund for Agricultural Development
IMF	International Monetary Fund
IPMS	Improving Productivity and Market Success
LM	Logic Model
M&E	Monitoring and Evaluation
MDG	Millennium Development Goal
MSH	Management Sciences for Health
NGO	Non-governmental Organization
OECD	Organization for Economic Co-operation and Development
OoARD	Office of Agriculture and Rural Development
PF	Performance Framework
PI	Performance Indicator
PLWs	Pilot Learning <i>Woredas</i> (Districts)
PM&E	Participatory Monitoring and Evaluation
PMF	Performance Measurement Framework
PRA	Participatory Rural Appraisal
RALC	Regional Advisory and Learning Committee
RBM	Results Based Management
RBM&E	Results Based Monitoring and Evaluation
RC	Result Chain
UNDP	United Nations Development Program
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WALC	<i>Woreda</i> Advisory and Learning Committee

## Acknowledgements

This guide has benefited from the contributions of a number of colleagues. We would like to express our special gratitude to Jerry Rogers (Results-Based Management Consultant) for his valuable comments on the draft guide and in helping us to incorporate recent developments in the result chain concept. We would also like to greatly thank Dr Ponniah Anandajayasekeram (Manager, ILRI CaSt) for his intensive and constructive comments throughout the draft which really helped us improve the guide. Dirk Hoekstra, IPMS Project Manager, has provided us with useful comments, an overall support and encouragement throughout the process of writing the guide. We would like to thank him immensely for that. We also thank all the experts of the Federal MoARD, and the Regional and Zonal BoARD of Tigray and Amhara who participated in the training workshops we organized on Result-based Monitoring and Evaluation. They have provided us with useful comments and feedbacks.



# 1 Introduction

Agricultural development is crucial to reduce poverty and achieve sustainable food security in many agriculture-based developing countries. The success and speed with which agricultural development is achieved depends in part on the performance of organizations working to promote development of the agricultural sector. Both governmental and non-governmental organizations working in agricultural development are increasingly concerned with the need to assess and understand their performance, and to improve relevance, effectiveness and efficiency of projects/programs/policies.

Currently, there is a significant emphasis on achieving results (outcomes) and on the need to demonstrate performance. The questions that are being asked by stakeholders have become: Are development initiatives making a difference in people's livelihoods? How will governments know whether they have made progress in bringing changes in people's livelihoods? Have projects/programs/policies led to the desired results (outcomes)? How can we tell success from failure? Do we know our starting points (baselines) in relation to how far we want to go? These are the kinds of concerns and questions being raised by development practitioners, other internal and external stakeholders, and governments across the globe are struggling with ways of addressing and answering them.

The increased level of emphasis given to results (outcomes), as opposed to activities and outputs, has also brought about some major changes in the focus, approach and application of monitoring and evaluation (M&E) systems. As the focus of management changes from activities to results, the focus of M&E also changes from the traditional M&E system, which focuses on assessing inputs and implementation processes (progress monitoring) to results-based M&E (RBM&E) system, which emphasizes assessment of the contributions of intervention to development outcomes.

In general, RBM&E deals with the measurement and assessment of performance in order to more effectively produce results (outcomes) so as to ensure that efforts are translated into changes in the lives of beneficiaries and their environment. The systems of RBM&E are means to measure the goods and services (outputs) that the organizations provide and to measure the extent to which the outputs are used by beneficiaries and how the living conditions of beneficiaries and their natural environment are changing as a result (Mackay 2007).

Understanding of the causes of good and poor performance is vital to improve the management of organizations, projects/programs and policies. RBM&E can provide vital information (and sometimes unique information) about the performance of projects/programs/policies. It can also be used to analyse what works and what does not, and the reasons why. As a result, a number of organizations are striving to improve their performance by developing systems to measure their performance. Therefore, RBM&E is being increasingly recognized as indispensable management function that helps organizations to improve performance and achieve desired results.

Among other things, the accountability and learning functions of RBM&E contribute to the increasing emphasis in developing strong RBM&E systems. First, RBM&E helps to build greater transparency and accountability with regard to the use of organizational resources. Stakeholders are no more interested only in resource use, organizational activities and outputs; they are now more interested in achieving results (outcomes) in terms of better access to services and improved livelihoods of beneficiaries. Budget shortages and growing expectations from clients force organizations to provide more services with higher standards of quality. Civil society and parliaments are also putting accountability pressures on governments to publicly report and explain their performance. Similarly, aid agencies are increasingly

requiring evidence of the results of their aid spending, pressurizing organizations to operate in most cost effective ways.

Secondly, learning is also facilitated by RBM&E. Information generated through RBM&E provides managers/staff with a clearer basis for decision-making. Future planning and implementation of projects/programs/policies is improved when guided by lessons learned from past experiences. RBM&E can help organizations to extract relevant information that can subsequently be used as the basis for planning, projects/programs/policies fine-tuning and reorientation. Without an effective RBM&E, it would be difficult to determine if an organization's work is going in the right direction or not, whether progress and success are being achieved or how future efforts might be improved.

However, it must be noted that RBM&E has no inherent value unless the RBM&E information is used to assess performance and improve organizational management, and the development and implementation of projects/programs/policies. RBM&E should be used as a learning tool and as an input to improve performance. The utilization of RBM&E information is critical to the institutionalization, performance and sustainability of the system.

In Ethiopia, there is an increasing emphasis among public agricultural organizations to improve performance and demonstrate results (outcomes). However, the use of RBM&E information to assess performance of projects/programs/policies has been severely limited in the country. Learning from past experiences in order to perform better in future has not been widely utilized by public and private organizations dealing with agricultural development. The main reasons are the low level of institutionalized RBM&E system and associated limited capacity to undertake RBM&E.

The purpose of this guide, therefore, is to contribute to the development of RBM&E capacity and to facilitate its institutionalization in organizations dealing with agricultural development. The target audiences of the guide include the staff in planning, monitoring and evaluation departments/units of public organizations and non-governmental organizations dealing with agricultural development at federal, regional, zonal or district levels. Staff of the agricultural research and higher learning institutes may also find the guide useful. It is assumed that users of the guide would have some basic knowledge of project/program/policy planning and implementation.

The guide is based on an extensive review of M&E literature and the experiences of the RBM&E activities of the IPMS (Improving Productivity and Market Success) of Ethiopian farmers project.<sup>1</sup> As part of its overall approach to market-oriented agricultural development, the IPMS project is working to facilitate the use and institutionalization of RBM&E system.

The guide is organized as follows. Section two deals with basic concepts of RBM&E. Section three presents the relationships between the concepts and practices of M&E. Section four deals with the concepts and applications of participatory monitoring and evaluation. Sections five and six present the practices and processes of the selection of results to monitor and evaluate, and the selection of key performance indicators, respectively. Section seven discusses the methods of setting baseline data and targets, and section eight deals with data collection and analysis. While section nine deals with reporting and using M&E information, section ten discusses issues, approaches and requirements for institutionalizing and sustaining the RBM&E system.

---

1. For more information about the IPMS project, visit the project website at [www.ipms-ethiopia.org](http://www.ipms-ethiopia.org).

## 2 Results-based monitoring and evaluation

### 2.1 Results-based management

Clearly, the increasing emphasis on results influenced the management of organizations and interventions, and necessitates the adoption of the Result-Based Management (RBM) approach. RBM is a participatory and team-based management approach that seeks to focus an organization's efforts and resources on expected results, improving effectiveness and sustainability of projects/programs/policies, and to improve transparency and accountability.<sup>2</sup>

RBM provides a coherent framework for strategic planning and management by improving learning and accountability. It is also a broad management strategy aimed at achieving important changes in the way agencies operate, with improving performance and achieving results as the central orientation, by defining realistic expected results, monitoring progress toward the achievement of expected results, integrating lessons learned into management decisions and reporting on performance. Therefore, RBM takes the focus away from activities to results. In general, RBM involves identifying project/program/policy beneficiaries, designing projects/programs/policies to meet their needs, defining realistic expected results, identifying and managing risks, monitoring progress towards results and resource consumed, increasing knowledge by learning lessons, incorporating lessons learned into management decisions, and reporting on the results achieved and resources involved to relevant stakeholders (CIDA 2009).

### 2.2 Results chain or logical model

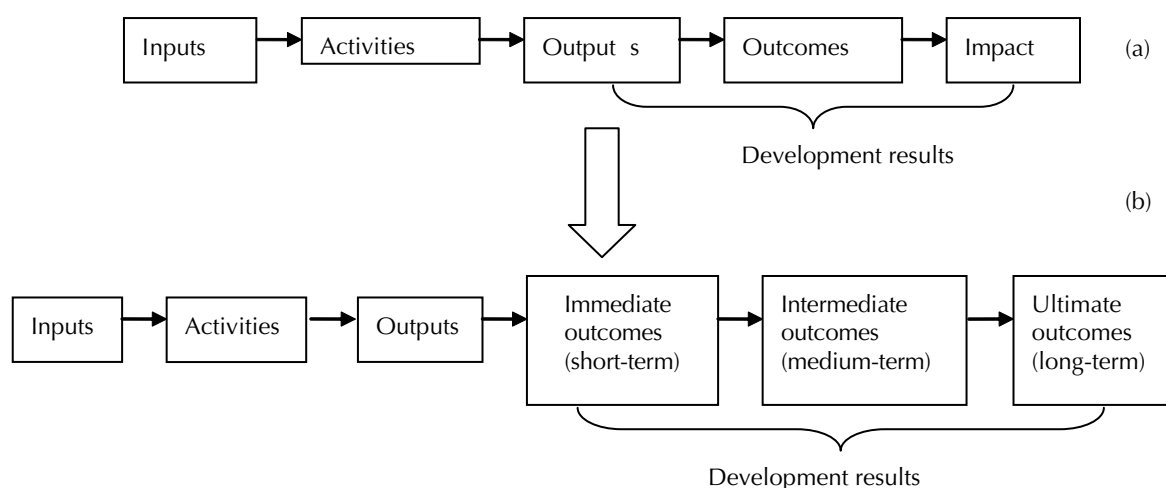
The basis for RBM is results chain (RC) or logic model (LM). RC is an illustration of the causal or logical relationships between the inputs, activities, outputs, and outcomes of a given project/program/policy (CIDA 1996). However, the elements included in the RC and what constitutes a result varies in different organizations and has changed over time. For example, in 2008, CIDA has made a slight change in RC from the one developed in 1996. Below we present differences between 1996 and 2008 CIDA RC and that developed by the Organization for Economic Cooperation and Development (OECD) Development Assistance Cooperation (DAC) in 2002.

#### *RC concept developed by CIDA in 1996 and 2008*

Figure 1 below shows the concepts and logical relationships between the different concepts of RC as developed by CIDA. The results chain in 1996 has five levels: inputs, activities, outputs, outcomes and impact (Figure 1a). However, the result chain in 2008 has six levels: inputs, activities, outputs, immediate outcomes, intermediate outcomes and ultimate outcomes (Figure 1b). Each level represents a step in the casual logic of a project/program/policy.

---

2. CIDA (2009) described RBM as 'a life-cycle approach to management that integrates strategy, people, resources, processes, and measurements to improve decision-making, transparency, and accountability. The approach focuses on achieving outcomes, implementing performance measurement, learning, adapting, as well as reporting performance.' Similarly, UNDP (2002) described RBM as 'a management strategy or approach by which an organization ensures that its processes, products and services contribute to the achievement of clearly stated results.'



Source: CIDA (2008a).

**Figure 1.** Results chain of CIDA 1996 (a) and of CIDA 2008 (b).

The slight difference in the 1996 and 2008 CIDA RC concepts is the division of outputs into outputs and immediate outcomes in 2008. Outputs in 1996 focus on changing knowledge, awareness, understanding and skills and can also be the direct products, goods or services from the activities. However, in 2008 outputs focuses on direct products or services stemming from the activities while immediate outcomes are usually at the level of an increase in awareness/skills, access to services etc. In addition, outcomes and impact are renamed as intermediate outcomes and ultimate outcomes, respectively.

Below we give description of the RC concepts of CIDA 2008.

The cause and effect linkages can be expressed by 'if...then' statements, representing the internal logic of the project/program/policy. For example, 'if' the activities are accomplished as planned, 'then' outputs are achieved; 'if' outputs are achieved as expected, 'then' we should achieve the immediate outcome, and; 'if' the immediate outcomes are achieved as expected, 'then' we should achieve the intermediate outcome and so on. While the first three levels (input, activity and output) represent the 'how' of an intervention, the last three levels (the immediate, intermediate and ultimate outcomes) address the actual changes (the developmental results) that take place. Below we give definitions of elements of the results chain as given by CIDA (2009) and examples from the IPMS project.

### *Inputs*

Inputs are financial, human, material, and information resources used to produce outputs through activities in order to achieve outcomes.

IPMS examples: Funds (money), time, human and other physical or material resources required to undertake activities to generate output to achieve market oriented commodity development.

### *Activities*

Activities are actions taken or work performed, through which inputs are mobilized to produce outputs.

IPMS examples: Giving training to farmers, identifying and promoting appropriate technologies, processes and institutional innovations for participatory market-led development programs in Pilot Learning *Woredas* (PLWs).

## Outputs

Outputs are direct products or services stemming from the activities of an organization, project/program/policy.

IPMS examples: Training completed, appropriate technologies, processes, and institutional innovations identified and promoted, research on market chains completed.

## Immediate (short-term) outcomes

Immediate outcomes are changes that are directly attributable to the outputs of an organization, project/program/policy. In terms of time frame and level, these are short-term outcomes, and are usually at the level of an increase in awareness/skills, access to services etc. among beneficiaries.

IPMS example: Increased knowledge, awareness, understanding and skills of staff of public organizations on approaches to market oriented agricultural development.

## Intermediate (medium-term) outcomes

Intermediate outcomes are changes that are expected to logically occur once one or more immediate outcomes have been achieved. In terms of time frame and level, these are medium term outcomes, which are usually achieved by the end of a project/program and usually occur at the change of behaviour/practice level among beneficiaries.

IPMS examples: Increased usage of knowledge gained from trainings, adoption of appropriate technologies, innovative input supply, output marketing, and financial services in order to improve agricultural productivity and market success in the PLWs.

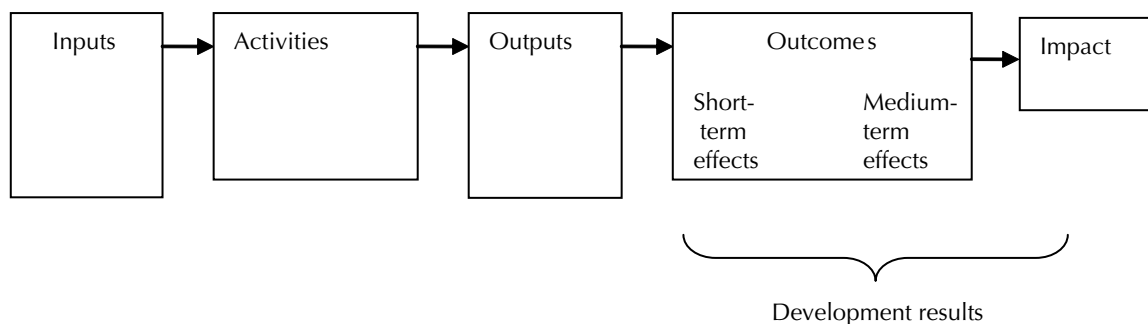
## Ultimate (long-term) outcome

Ultimate outcomes are the highest level changes that can be reasonably attributed to an organization, project/program/policy in a causal manner, and is the consequence of one or more intermediate outcomes. The ultimate outcome usually represents the *raison d'être* of an organization, project/program/policy, and takes the form of sustainable change of state among beneficiaries.

IPMS example: Improved agricultural productivity and production within functional market-oriented agricultural production systems in and beyond the PLWs.

## RC concepts developed by OECD DAC in 2002

Figure 2 below shows the concepts and logical relationships between the different concepts of RC as developed by OECD DAC. The DAC concept distinguishes between outcomes and impact. Outcomes are also divided into short-term and medium-term.



Source: OECD DAC (2002).

**Figure 2.** Results chain.

Below we give definitions of elements of the results chain as given by OECD DAC (2002).

### *Inputs*

Inputs are the financial, human, and material resources used for development interventions.

### *Activities*

Actions taken or work performed through which inputs, such as funds, technical assistance, and other types of resources are mobilized to produce specific outputs.

### *Outputs*

Outputs are the products, capital goods and services that result from a development intervention; may also include changes resulting from the intervention that are relevant to the achievement of outcomes.

### *Outcomes*

Outcomes are the likely or achieved short-term and medium-term effects of an intervention's outputs.

### *Impact*

Impacts are positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.

## Comparisons of the RC concepts given by OECD DAC and CIDA

In general, as can be seen in Figures 1 and 2 above, the result chain followed by CIDA and OECD are more or less the same. However, there are slight differences in the terminologies used to describe components of development results. Specifically, the definition that CIDA uses for immediate outcome (short-term) is what OECD DAC refers to as short-term effect outcomes. The definition that CIDA uses for an intermediate outcome (medium-term) is what OECD DAC refers to as medium-term effect outcomes. However, in OECD DAC, short and medium term effect outcomes are presented together as outcome rather than separately as is in CIDA. In addition, the definition that CIDA uses for the term ultimate outcome (long-term) is what OECD DAC refers to as impact. In this guide we adopt the RC concept developed by CIDA.

## 2.3 Performance framework

Performance refers to the extent to which a development intervention operates according to specific criteria/standards/guidelines or achieves results in accordance with stated goals or plans (OECD 2002). A well-performing organization or project/program/policy is one that is providing, in the most cost effective manner possible, expected results that continue to be relevant, without causing unintended negative consequences.

Performance measurement is part of results-based management, and is the basis for RBM&E. Performance measurement refers to the measurement and comparison of the performance of development interventions against stated goals (OECD 2002). Therefore, performance measurement is the ongoing monitoring and reporting of accomplishments, particularly progress towards pre-established goals and targets. Performance measurement is concerned with measuring both implementation progress (implementation measurement) and results achieved (result measurement) (DAC 2000). It may address whether or not project inputs and activities are in compliance with design budgets, work plans and schedules, it may also measure the direct products and services delivered by a project/program/policy

(outputs), and/or the change in access to services (immediate outcomes), the utilization of the products and services (intermediate outcomes), or the effect of the outcomes on people's livelihoods and the environment (ultimate outcome). Performance measurement helps to answer the question of whether development intervention is achieving the results that stakeholders expect. It also helps organizations to learn from the answer to this question and manage more effectively.

A performance framework (PF) is a RBM tool that depicts the concept of a project/program/policy. It identifies resources, reach, goal and purpose of the intervention as well as the cause and effect relationships among activities, outputs and a chain of results (CIDA 1999). Below we present a brief description of the concepts included in a PF and illustrate with examples from the IPMS project:

#### *Goal:*

Goal is the higher stated objective to which a development intervention is intended to contribute (OECD 2002). It is a statement of desired state where a need or a problem is addressed and it is related to the highest order result to which the project/program or organization contributes. It expresses the benefits to the target groups or beneficiaries, although it cannot normally be achieved by one project/program or organization alone.

IPMS example: To contribute to improved agricultural productivity and production through market-oriented agricultural development, as a means for achieving improved and sustainable livelihoods for the rural population.

#### *Purpose:*

Purpose is the publicly stated objective of the development project/program/policy (OECD 2002). It is a measurable statement of the outcome of an organization or project/program within a given period. It is related to what an organization, project/program is expected to achieve in terms of outcome.

IPMS example: To strengthen the effectiveness of the Government's effort to transform agricultural productivity and production, and rural development in Ethiopia.

#### *Resources:*

Resources refer to human, organizational and intellectual and physical/material inputs that are directly or indirectly invested by an organization or project/program (Montague as cited in CIDA 1999). It includes the amount of time, money and/or energy exerted and the type of resources used.

IPMS example: Finance, material inputs for commodity development, staff, vehicles and equipment used to implement interventions for market oriented agricultural development.

#### *Reach:*

Reach refers to the breadth and depth of influence over which the organization or project/program/policy wishes to spread its resources (Montague as cited in CIDA 1999). While physical (spatial) reach is one dimension, it also includes the type of groups the intervention wishes to affect.

IPMS example: Rural farmers (women and men), public and private organizations that support the agricultural sector, and 10 pilot learning districts in 4 regions etc.

#### *Outputs:*

Outputs are direct products or services stemming from the activities of an organization and project/program/policy.

IPMS example: see section above.

*Results:*

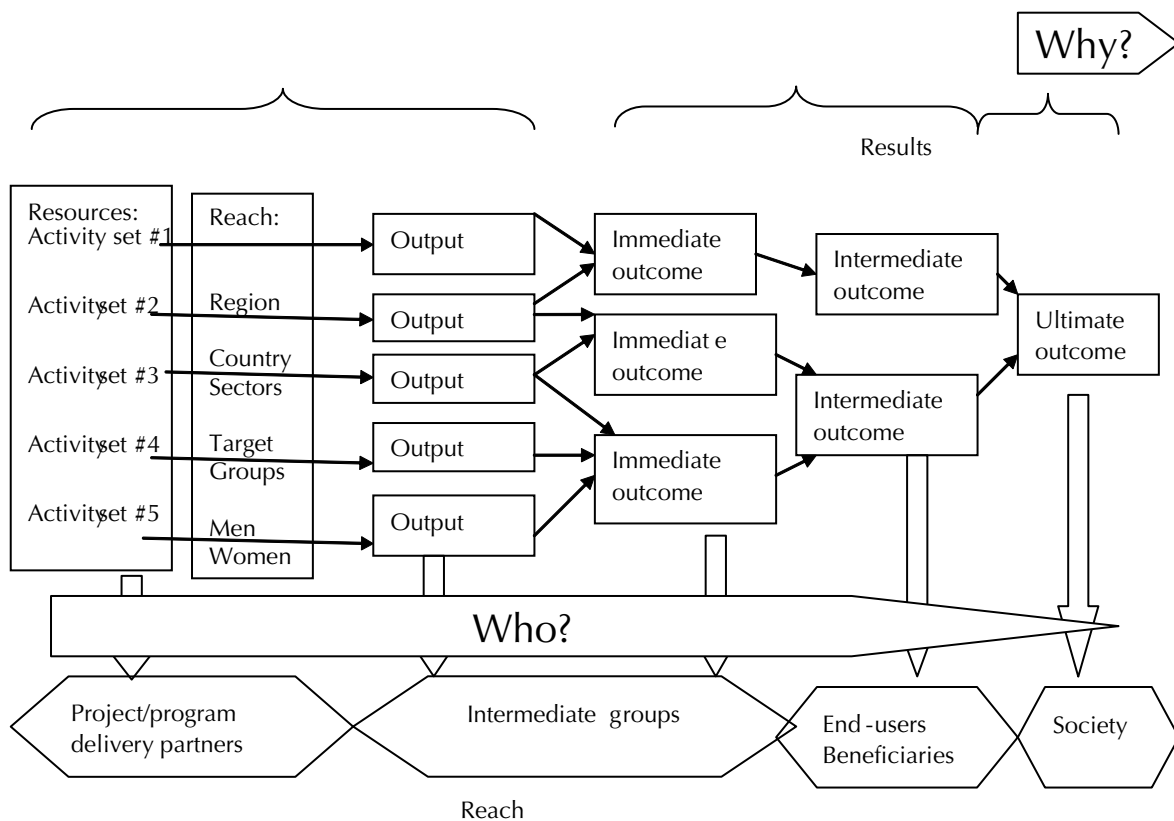
Results are describable or measurable changes that are derived from a cause and effect relationship (CIDA 2009). Results are outcomes and further classified into immediate (short-term), intermediate (medium-term) and ultimate (long-term) outcomes. A result statement gives information about what a project/program/policy is expected to achieve or contributes to.

IPMS example: see section above.

*Activities:*

Activities are actions taken or work performed through which inputs are mobilized to produce outputs.

Figure 3 shows a schematic representation of a PF, including goal and purpose.



Sources: Adapted from CIDA (1999, 2009).

**Figure 3.** Building a performance framework.

**Goal:** The strategic objective to which an intervention is intended to make a contribution.

**Purpose:** The publicly stated objective of the development project/program/policy which addresses the priority needs of the intended beneficiaries.

## 2.4 Performance measurement framework

As discussed above, measuring performance is one of the major components of the RBM approach. Therefore, it is important to establish a structured plan for data collection, analysis, use and dissemination of performance information. In this regard a performance measurement framework (PMF) is an important tool that is used to structure basic information needed for performance measurement. A PMF is a plan



to systematically collect relevant data over the lifetime of an intervention to assess and demonstrate progress made in achieving expected results (CIDA 2009). A PMF is presented in a table form to document the major elements of the monitoring system and ensures that performance information is collected in a regular basis. It may have about eight columns: expected results, performance indicators, baseline data, targets, data sources, data collection methods, frequency of data collection and responsible actors for data collection and analysis. Figure 4 below shows a stylized PMF.

Expected results	Performance indicators	Baseline data	Targets	Data sources	Data collection methods	Frequency of data collection	Responsibility
Ultimate outcome							
Intermediate outcomes							
Immediate outcomes							
Outputs							

Source: Adapted from CIDA (2009).

**Figure 4.** Performance measurement framework.

Building a PMF starts with the expected results column which provides the outputs, immediate outcomes, intermediate outcomes and ultimate outcome (see section five for more information about selecting result expectations). The next step is establishing the performance indicators for the ultimate, intermediate and immediate outcomes and the outputs. Performance indicators (PI) are variables to measure results achieved and are used to monitor project/program/policy performance (see section six for details in how to select performance indicators). Then sources of data and data collection methods are identified and recorded in the PF. Data sources refer to the individuals, organizations or documents from which data about indicators will be obtained. Examples of data sources include beneficiaries, stakeholders, government or donor agency documents and/or statistical reports. Data collection methods refer to approaches and techniques of how data about indicators are collected (refer section eight for more information on data collection and analysis).

The frequency of data collection and the actors responsible for data collection and analysis are then identified and recorded. Frequency refers to the timing of data collection and gives information about how often information is collected for each indicator. Information can be collected annually, bi-annually etc. Responsibility refers to who is responsible for collecting and/or validating the data. It can also include responsibility for data analysis and reporting. Some examples of actors who can be responsible for data collection/validation include local professionals, partner organizations, organizational staff or consultants.

Finally, baseline data will be filled and realistic targets are established for each indicator. Baseline data is the set of conditions existing at the outset of a project/program/policy. A target specifies a particular value for a performance indicator to be achieved by a specific date in the future (refer section seven for more information about establishment of baseline data and targets). Hence, a completed PMF documents the major elements of the monitoring system and ensures that comparable performance information is collected on a regular and timely basis.

## 2.5 Implementation-focused vs. results-based M&E

Implementation is concerned about how well outputs are achieved using available inputs and activities. Hence, recently, based on focus, distinctions have been made between implementation-focused and

results-based M&E systems (RBM&E). According to DAC (2000), implementation-focused M&E is the frequent, ongoing documentation of data in operations such as tracking of funds and other inputs and processes. On the other hand, RBM&E focuses on the actual achievement of results. Implementation-focused M&E systems are designed to address issues of compliance with plans, such as answering questions like ‘did we procure the needed inputs?’, ‘did we implement the planned activities?’, ‘did we do it as planned?’, ‘did we achieve the planned outputs?’ (Kusek and Rist 2004). Implementation-focused approach emphasizes monitoring and assessing how well a project/program/policy is implemented relative to plans. Such M&E system fails to provide decision-makers and stakeholders with an understanding of the success or failure of the project/program/policy with regard to meeting intermediate and long term results.

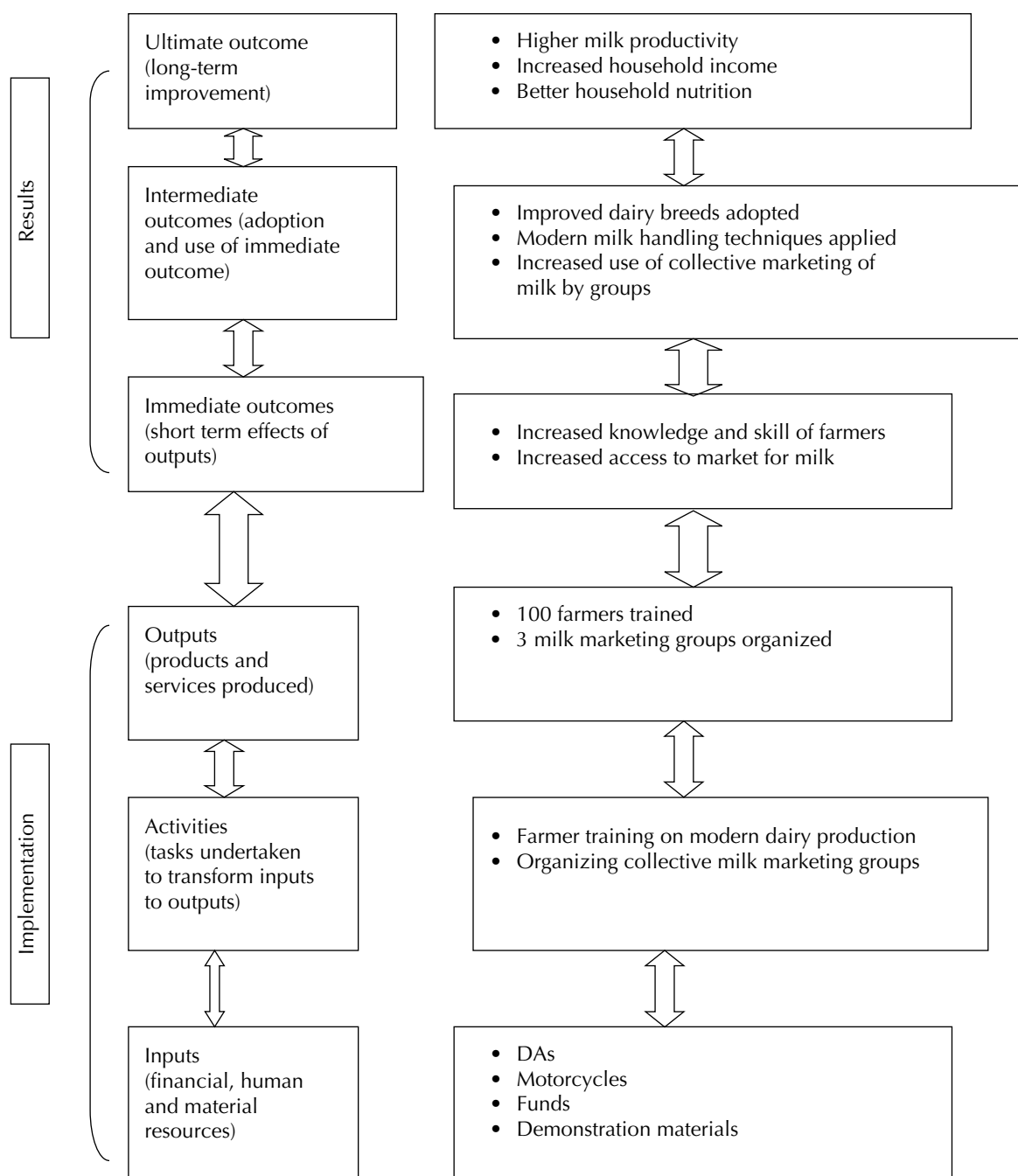
RBM&E aims at expanding the implementation M&E function to include results (outcomes) explicitly. It is designed to answer the ‘so what?’ questions. RBM&E addresses questions such as ‘so what that activities have taken place?’, ‘so what that outputs have been produced?’. Hence, RBM&E provides feedback on different level of outcomes of intervention. RBM&E uses more qualitative and quantitative information on the progress towards outcomes. It attempts to answer the following key questions (Kusek and Rist 2004): (1) What are the results of the intervention? (2) Are the results of the intervention being achieved? (3) Are there ways by which achievements could be articulated and measured?

Figure 5 shows how M&E system should include not only the implementation focus, but also a results focus, and how RBM&E system builds on the implementation-focused systems. The left hand column shows the RC and the right hand column shows examples of the components of the RC.

## 2.6 Uses of M&E information

In M&E system, information on progress should be collected for all result levels, at different time frames, and for different stakeholder needs. In addition to providing guidance to improve the performance of interventions and organizations, RBM&E systems has several additional advantages. A sound M&E system can be a source of knowledge, enabling organizations to develop a knowledge base of the type of projects/programs/policies that worked and did not work, and why, thus promoting organizational learning. M&E systems are also instrumental in promoting transparency and accountability within organizations. Hence, the major objectives or applications are:

1. To monitor and evaluate a project/program/policy  
Information can be collected and analysed at each or all levels on a continuous basis so that the data can be used to provide timely and useful information to decision-makers and stakeholders. As such, M&E should be conducted throughout the design and implementation cycle of projects/programs/policies, as well as after completion.
2. Accumulation of knowledge  
Good M&E system helps governments and organizations to develop knowledge base of the types of projects/programs/policies that have worked and did not work, and why. In addition, M&E systems provide continuous feedback thus promoting organizational learning.



Sources: Adapted from Kusek and Rist (2004); CIDA (2008b).

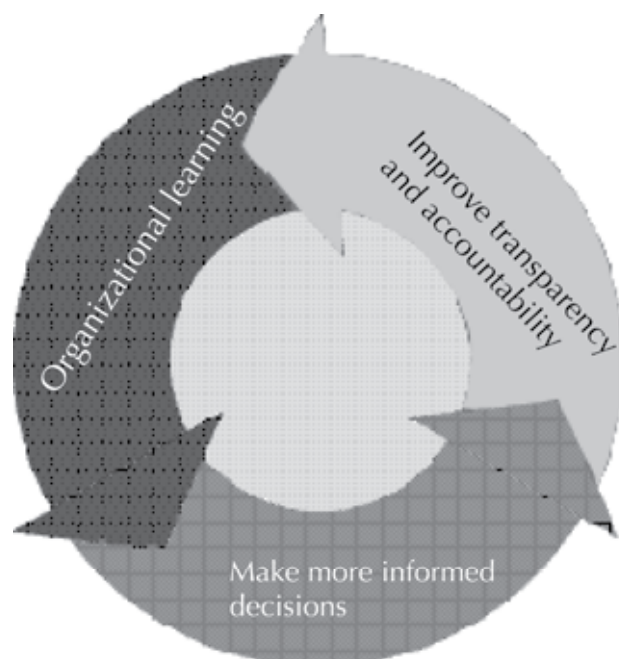
**Figure 5.** Integrating results-based M&E with implementation-focused M&E.

### 3. Transparency and accountability

M&E systems can be used to promote greater transparency and accountability within organizations and governments. Internally it can serve as a management tool to take corrective action and help future planning and effective resource allocation. This would help external and internal stakeholders have a clearer idea of the status of projects/programs/policies. In this regard, demonstrating positive results can help get greater political and popular support.

The three main objectives of M&E (to enhance organizational learning, ensure informed decision-making, and support transparency and accountability) are linked to one another in a continuous process (Figure 6). Knowing which activities contribute (not contribute) to achieving goals gives lessons

to staff and managers. This learning from the past contributes to more informed decision-making. Better decisions lead to greater accountability to stakeholders.



Source: Adapted from UNDP (2002).

**Figure 6.** *Links between the major objectives of M&E.*

## 3 Relationships between the concepts and practices of monitoring and evaluation

### 3.1 Monitoring information to support result-based management (RBM)

Monitoring involves the collection of routine data that measures progress towards achieving project/program/policy objectives. It is used to track changes in the intervention performance over time. Its purpose is to permit stakeholders to make informed decisions regarding the effectiveness of the intervention and the efficient use of resources. The word monitoring is defined differently by different organizations. However, the basic idea of the concept remains the same for most of the definitions. For example, OECD defines monitoring as:

‘A continuous function that uses systematic collection of data on specific indicators to provide management and main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds’ (OECD 2002).

Similarly, UNDP defines monitoring as:

‘A continuous function that aims primarily to provide management and main stakeholders of an ongoing project [intervention] with early indications of progress, or lack thereof, in the achievement of results’ (UNDP 2002).

Hence, the essence of monitoring is the planned and organized collection of data that will help us answer questions about progress (or lack of it) in an organization or, project/program relative to plans. It provides regular follow-up and oversight of an organization’s or project/program work status in terms of input delivery, activities, targeted outputs and outcomes. The monitoring system provides ongoing information based on selected indicators, on the direction of change, the pace of change, and the magnitude of change achieved as a result of implementing a project/program/policy, while also identifying unanticipated changes. Through such routine data gathering, analysis and reporting, monitoring provides management, staff and other stakeholders with information on whether progress is being made towards achieving results. In this regard, the objectives of monitoring can be categorized as follows (Anandajayasekera et al. 2004):

- Record inputs, activities and outputs
- Identify deviations from work plans
- Identify constraints/bottlenecks
- Assess overall efficiency
- Assess overall effectiveness
- Assess suitability of new methods and technologies under testing
- Assess the long-term impact
- Learn from achievements and mistakes
- Increase capacity to perform better
- Take corrective actions
- Share progress and results with others

Monitoring can be progress or process monitoring. Process monitoring is different from progress monitoring in that the latter focuses on physical, financial and logistic aspects of projects/programs,

while process monitoring deals with critical processes which are directly related to the projects/ programs objectives. For example, progress monitoring looks at the number of training sessions held, or the percentage of work completed; while process monitoring evaluates the quality of training or the level of community involvement. Elliott et al. (2005) defined process monitoring as 'a set of activities of consciously selecting processes, selectively and systematically observing them so as to compare them with others and communicating on that in order to learn how to steer and shape the processes.' Process monitoring helps to learn from own experience and adapt to improve effectiveness over time. An ideal M&E system contains elements from both progress and process monitoring.

Monitoring is an internal project/program management tool, used to systematically and critically observe progress in order to manage activities and adapt them to changing implementation conditions. Integrating monitoring with the implementation process has the following advantages:

- increases the accuracy of the collected data and information,
- reduces the cost of data and information collection,
- increases the focus of participating implementers,
- reduces the time lag for management corrections.

Being a continuous function of an organization, monitoring is usually carried out by organization's own staff or together with relevant stakeholders. In order to do effective monitoring, besides qualified personnel, an organization needs adequate planning, baseline data, indicators of performance, and practical implementation mechanisms (e.g. field visits, stakeholders meetings, documentation of activities, regular reporting).

The key steps followed in the monitoring process include:

- Recording data and information on key indicators, mainly from sources existing at the organization, such as existing financial records, supply procurement and disbursement books
- Collection of primary data
- Computerization of data
- Analysis performed at each functional level of management
- Regular reporting, either in written form or orally, such as quarterly, semi-annually or annually
- Manual or computerized storage of data and information

### 3.1.1 Implementation monitoring vs. results monitoring

Monitoring can be classified into two: implementation monitoring and results monitoring. Both types of monitoring are important in tracking results and they are complimentary.

#### *Implementation monitoring*

Implementation monitoring tracks the means and strategies (i.e. inputs, activities and outputs stipulated in work plans) used to achieve an outcome. The means and strategies are backed up by budgetary resources, staffing and activity planning. Annual work plans are the means and strategies that are used to effectively conduct activities and achieve outputs, and ultimately outcomes. Every target must be viewed as an intermediate effort on the way to achieving an outcome. Hence, means and strategies should be implemented to help achieve targets.

#### *Results monitoring*

Results monitoring is concerned with how outputs are translated into different levels of outcomes. However, it must be stressed that the interaction between means and strategies (inputs, activities and

outputs) and outcome targets is crucial in achieving the overall development goal of an intervention. Hence, while implementation monitoring is concerned with how outputs are achieved using inputs and activities, results monitoring is concerned with the alignment of the outputs with outcomes.

## 3.2 Evaluative information to support results-based management systems

Evaluation measures how well an intervention has met expected objectives and/or the extent to which changes in results can be attributed to the intervention. As for monitoring, there are different definitions of the concept of evaluation, although the essence of the concept remains the same for most of the definitions. For example, OECD (2002) defined evaluation as:

‘...the systematic and objective assessment of an ongoing or completed project, program or policy including its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability.’

Similarly UNDP (2002) defined evaluation as:

‘...a selective exercise that attempts to systematically and objectively assess progress towards and the achievement of an outcome. Evaluation is not a one-time event, but an exercise involving assessments of differing scope and depth carried out at several points in time in response to evolving needs for evaluative knowledge and learning during the effort to achieve an outcome. All evaluations, even project evaluations that assess relevance, performance and other criteria need to be linked to outcomes as opposed to only implementation or immediate outputs.’

Hence, the essence of evaluation is using monitoring and other information to make judgements about an intervention. It is important to note that the function of evaluation in M&E system expands and moves beyond the traditional after-the-fact approach to evaluation. Evaluation should not be restricted to assessing causes and changes after an intervention is over. The after-the-fact approach does not provide information that can be used as an input in an ongoing intervention aimed at achieving results. Hence, good evaluative information is needed throughout the life cycle of a development intervention.

Evaluations often document and explain the causes as to why activities succeed or fail. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process. Evaluation provides managers with information regarding performance and can provide signs of strengths and weaknesses, and therefore, enable managers to improve future planning, delivery of services and decision-making. Such documentation can help in making future activities more relevant and effective. It also assists managers, staff and other stockholders to determine in a systematic and objective manner the relevance, effectiveness, efficiency, impact and sustainability of activities and their results. Evaluations also help to generate detailed information about implementation process and results. Such information can be used for public relations, promotion of services in the community, as well as identifying possibilities for replication of activities.

### 3.2.1 Types of evaluations

Evaluations can be classified based on (1) who conducts the evaluation, (2) when they occur in the intervention cycle, or (3) based on the types of questions they are expected to answer. Based on who conducts the evaluation, evaluations are classified into internal, external and collaborative/joint evaluations. Based on when they are conducted, evaluations are categorized into *ex-ante* (before the

intervention), ongoing (during the intervention), and *ex-post* (immediately after the intervention or several years after the intervention) (Anandajayasekeram et al. 2004). Based on the type of questions that an evaluation is expected to answer, evaluations are classified into performance logic chain assessment evaluation, pre-implementation assessment evaluation, process implementation evaluation, impact evaluation and meta evaluation (Kusek and Rist 2004). Below we give brief descriptions of the types of evaluations.

Based on who conducts the evaluation:

#### *Internal evaluation*

In internal evaluation, sometimes called self-evaluation, a unit and/or individuals reporting to the management of the donor, partner or implementing organization conduct the evaluation. The advantage of using internal evaluator is that insiders know the organization and therefore may be able to interpret the results better than an external body. The disadvantage of using internal evaluator is that internal evaluator may avoid negative conclusions. In other words, strengths and weaknesses might not be interpreted fairly when data and results are analysed by internal staff members.

#### *External evaluation*

This is a type of evaluation in which the evaluation of a development intervention is conducted by entities and/or individuals outside the implementing or donor agency. Many organizations may not have the resources to carry out the ideal evaluation. In such cases external evaluation consultant is recruited to lead the evaluation process. An external evaluator may be more objective, free from organizational bias and may contribute fresh perspectives.

#### *Joint evaluation*

In joint evaluation different implementing and donor agencies as well as partners participate in the evaluation. The degree of 'jointness' may vary depending on the extent to which individual partners cooperate in the evaluation process, contribute resources for the evaluation and combine their evaluation reporting. Joint evaluation can help overcome the problem of attribution problems in assessing the effectiveness of programs and strategies and the complementarity of efforts supported by different partners etc.

Based on when they are conducted:

#### *Ex-ante evaluation:*

An *ex-ante* evaluation is made to assess the potential impact of a project, program or policy intervention before implementation. *Ex-ante* evaluation is a process that supports the preparation of proposals for new interventions. Its purpose is to gather information and carry out analyses that help to ensure that the objectives can be met and that the method used is cost-effective. It is done to estimate costs and benefits and assesses the potential impact of an intervention before it is implemented. *Ex-ante* evaluations are done by peer or expert reviews using checklists, scoring models, or cost-benefit analysis.

*Ex-ante* evaluation can provide an idea of what range of impact to expect after the project/program/policy is implemented. It can also assist in setting up an appropriate M&E system for *ex-post* impact assessment. Moreover *ex-ante* evaluation methods can be used to identify how the impacts would change if some parameters of the program were changed. *Ex-ante* evaluation is a tool for improving



the quality of new or renewed projects/programs and for providing information on the basis of which decision-makers can judge the value of a proposal. Therefore, it is important to start *ex-ante* evaluation work early on in the process when options for project/program formulation are still open.

#### *Ongoing evaluation:*

Ongoing evaluations review ongoing activities to provide guides for corrective implementation measures in order to achieve intended results better. As such, ongoing evaluation is conducted during the implementation stage. Periodic evaluation of ongoing interventions is conducted to analyse the use of resources, the quality of work, and the continuing relevance of the intervention. It is also used to review implementation progress and predict likely effects of interventions and highlight necessary adjustments in work design. Mid-term evaluation which is conducted at the middle of a project/program life, serve as a means of validating the results of initial assessments obtained from monitoring activities. Ongoing evaluations address problems associated with the day-to-day management of interventions and also can indicate the need for changes in project, program or policy objectives and targets.

#### *Ex-post evaluation:*

An *ex-post* evaluation assesses the interventions performance, quality, relevance, efficiency and impact immediately after implementation is completed. An *ex-post* evaluation is linked to an *ex-ante* evaluation, and is best conducted where a baseline has been originally defined, targets have been projected, and data has been collected on important indicators. Information collected through monitoring is also fundamental for the success of *ex-post* evaluation.

This kind of evaluation provides an overall assessment of the intervention's performance, cost effectiveness, its relevance to development goals, and acceptance of the results by end users and/or its impacts. *Ex-post* evaluation also assesses the extent to which an intervention has succeeded in meeting its objectives. Moreover, in addition to providing information about potential sustainability of an intervention, *ex-post* evaluation helps to obtain feedback from the target group. Most of the time *ex-post* evaluation is carried out through participatory meetings at the site with peers, farmers, extension staff, NGOs and other relevant stakeholders. Impact evaluation is a form of *ex-post* evaluation and attempts to determine the extent to which the intervention has contributed to the achievement of the larger development goal (see next section for more on impact evaluation).

### Based on the types of questions they are expected to answer

Different types of evaluations are required for different types of questions. It is important for decision-makers to know what type of evaluative information they need to have.

#### *Performance logic chain assessment evaluation*

The performance logic chain assessment evaluation is used to evaluate the strength and logic of the causal model underlying the project/program/policy. The causal model addresses the deployment and sequencing of the resources, activities or policy strategies that can be used to achieve a desired change in an existing situation. The purpose of a performance logic chain assessment evaluation is to avoid failure from a weak or inappropriate design of an intervention.

#### *Pre-implementation assessment evaluation*

As the name indicates, this type of evaluation is conducted just before the implementation phase. The pre-implementation assessment evaluation attempts to answer three fundamental questions that underlie an intervention: (1) Are the objectives well defined so that outcomes can be stated in measurable terms? (2) Is there a coherent and credible implementation plan that provides clear evidence of how implementation is to proceed, and how successful implementation can be distinguished from poor implementation? (3) Is the rationale for the deployment of resources clear and commensurate with the requirements for achieving the stated outcomes? The purpose of such evaluation is to ensure that failure is not programmed in from the beginning of implementation.

#### *Process implementation evaluation*

The process implementation evaluation focuses on the implementation details. Its defining theme is a focus on the intervention itself, i.e. its operation, activities, functions, performance, component parts, resources etc. (Rossi et al. 2004). Process implementation evaluation verifies what the intervention is and whether or not it is delivered as intended to the targeted recipient. It does not attempt to assess the effects of the development intervention on those recipients (such assessment is the concern of impact evaluation, which is discussed next). Process implementation evaluation involves assessment of the performance of an intervention with regard to service utilization and organization of the intervention. Assessing service utilization consists of examining the extent to which the intended target population receives the intended services. On the other hand assessing the organization of the intervention requires comparing the plan for what the intervention should be doing with what is actually done, especially with regard to providing services. Specifically it attempts to answer questions such as (1) What did or did not get implemented that were planned? (2) What congruence was there between what was intended to be implemented and what actually happened? (3) How appropriate and close to plan were the costs; the time requirements; the staffing capacity, and capability; the availability of required financial resources, facilities and staff; and political support? (4) What unanticipated and unintended outputs or outcomes emerged from the implementation phase? (5) Whether the intervention is reaching the target population (e.g. how many persons are receiving services? Are the intended targets receiving proper amount, type, and quality of services, are the target population aware of the project/program etc.)?

Decision-makers can use this information to determine whether they need to make any mid-course corrections to drive toward stated outcomes. This type of evaluation is similar to monitoring. The added value is that implementation is not just monitored (documented), but unanticipated outcomes are studied. Moreover, some intangible aspects of implementation such as political support, institutional appropriateness for change, management capability to achieve change can be addressed.

#### *Impact evaluation*

Impact evaluations are conducted some time after the completion of an intervention and their objective is to determine the effect of interventions on beneficiaries. Impact evaluation goes beyond the direct result of an intervention and tries to study the effects of the intervention on the ultimate users. However, such evaluations should distinguish between the contributions of the intervention under evaluation from the contributions made by other factors. Therefore, the key concepts in *ex-post* impact assessments are causality, attribution and incrementality. An impact evaluation attempts to find out the changes that occurred, and their attribution. This evaluation tries to determine what portions of the documented impacts are attributable to the intervention and what proportion to other events or conditions.

Impact evaluations attempt to answer the counterfactual question of 'what would have happened if the

intervention had not taken place?’ Therefore, all impact evaluations are inherently comparative, i.e. determining the impact of a development intervention requires comparing the condition of targets that have experienced an intervention with an estimate of what their condition would have been had they not experienced the intervention. This comparison is done by comparing changes for participants of the development intervention with those of equivalent persons who have experienced something else. There may be one or more groups of targets receiving ‘something else’, which may mean receiving alternative services or simply going untreated. The ‘equivalent’ targets for comparison may be selected in a variety of ways, or comparisons may be made between information about the outcome being examined and similar information from the same target taken at an earlier time (Rossi et al. 2004). Impact evaluation are better planned before the intervention begins, as it might help determine which units would receive the intervention and which will not, and establish baseline information on all units.

#### *Meta evaluation*

Meta evaluation is a term used for evaluations designed to aggregate findings from different evaluations (OECD 2002). Given a number of evaluations that are available on one or similar interventions, a meta-evaluation is used to establish the criteria and procedures for systematically looking across those existing evaluations to summarize trends and to generate confidence (or caution) in the cross-study findings. Meta evaluations attempt to answer questions like ‘what do we know at present on this issue and what is the level of confidence with which we know?’ Meta evaluations are quick way of learning from previously conducted evaluations.

### 3.2.2 Uses of evaluation information

There are several uses of evaluation information. Below we give a brief account of the types of uses.

#### *1. Selection among competing alternative strategies or resource allocation decisions*

Based on pilot experiences, evaluation information can be used to select best alternative strategies to address a development problem. Evaluation information can aid in resource allocation by identifying what projects/programs/policies have been successful in achieving desired outcomes. Similarly, evaluation information can be used to determine whether a pilot intervention should be scaled out, redesigned or dropped.

#### *2. Check on the effectiveness and efficiency of ongoing intervention*

Evaluation information can be used to assess the rationale or justification of an intervention strategy in order to provide feedback on whether the right things are done or not. Evaluation information is also used to assess the effectiveness of an intervention in achieving results, and the efficiency of resource use.

#### *3. Better understanding of problems*

Evaluation information helps achieve a better understanding of the development problems to be addressed or the identification of real priority development problems that an intervention should address. Evaluation information can also highlight emerging problems that are not yet widespread, but could deserve attention.

#### *4. Organizational learning*

Evaluation information can be used to build the knowledge base of an organization in terms of what projects/programs/policies work and do not work. In addition evaluation information can be used to document the factors and processes that contributed to the success or failure of an intervention. This information can effectively guide future planning.

### 3.2.3 Characteristics of quality evaluations

A good evaluation should have six characteristics: stakeholder involvement, impartiality, usefulness, technical adequacy, cost effectiveness and timely dissemination and feedback. We describe each characteristic briefly below.

#### *Stakeholder involvement*

Participation of stakeholders in the design and implementation of an evaluation facilitates trust in and ownership of the evaluation findings. Moreover, the chance that stakeholders would be willing to incorporate lessons in ongoing or future evaluations is greatly increased if stakeholders are involved in the evaluation process. Hence, participatory evaluation is a key feature of good evaluations.

#### *Impartiality*

It is critical to ensure that an evaluation process is free of political interference, biases and distortions. Reports of evaluation findings should include descriptions of the strengths and weaknesses of the evaluation process. Evaluation reports should be comprehensive in the sense that all relevant information is included.

#### *Usefulness*

Unless evaluation information is relevant, timely, and written in understandable form, its usefulness will be greatly diminished, since evaluations do not have inherent value. An evaluation should address the relevant questions that decision-makers need addressed in order to make decisions.

#### *Technical adequacy*

An evaluation process needs to collect data and information and analyse them in order to guide decisions. Data collection and analysis should follow relevant technical standards. Appropriate research and sampling designs, accurate wording of questionnaires and interview guides, appropriate qualitative and quantitative analysis methods should be used.

#### *Cost effectiveness*

An evaluation should be conducted in a cost effective manner. Gathering expensive data that will not be used should be avoided. The cost of the evaluation needs to be proportional to the overall cost of the intervention.

#### *Timely dissemination and feedback*

Evaluation findings should be shared in an appropriate, targeted and timely fashion. Lack of or delay in communication of evaluation information may lead to loss of trust, indifference or suspicion about the findings themselves.

## 3.3 Differences and complementarities of monitoring and evaluation

### 3.3.1 Differences between monitoring and evaluation information

Monitoring data does not provide the basis for attribution and causality for change, nor for evidence of how changes are being achieved. Monitoring cannot address the strengths and weaknesses in the design and implementation of project/program/policy. As a result, evaluation information is necessary to address these and other questions that remain unanswered by monitoring information.

Hence, monitoring and evaluation are two distinct functions, and yet complimentary to each other. Although both monitoring and evaluation can be done at project/program/policy levels, monitoring is concerned with checking on progress to determine if objectives are achieved or not; while evaluation is a more reflective process aimed at assessing an intervention and its results according to agreed criteria such as effectiveness, efficiency, quality, relevance, impact and sustainability. While monitoring gives information on where an intervention is at a given time or over time relative to targets, evaluation gives evidence of why targets are or are not achieved. As such, monitoring is descriptive in nature while evaluation attempts to address issues of causality and at times calls for value judgement.

In general, evaluation is much wider in scope than monitoring. It deals with making an assessment of overall achievements. An evaluation may address questions such as: Have we met the original objectives? Have we achieved the results we intended to achieve? How efficiently were the results achieved? Could we have achieved the output in another way, more effectively or more efficiently? What would have happened without the intervention? Monitoring usually leads to corrective action at the operational level, while evaluation leads to affirmation or modification of objectives, resources and processes.

### 3.3.2 Complementarity between monitoring and evaluation information

Monitoring serves a management function by trying to determine whether the material and financial resources are sufficient, whether the people in charge have the necessary technical and personal qualification, whether resource flows are consistent with the design, whether activities conform to work plans, and whether the work plan has been achieved and has produced the expected results. On the other hand evaluation complements monitoring by providing explanations of why there are or are not deviations between results and targets. In other words, when monitoring sends signals that implementation is deviating from plans (in terms of cost, services non-use, non-adoption of results, missed target populations etc.), then evaluation helps to clarify and explain the trends and conditions noted in the monitoring work.

The complementarities between monitoring and evaluation can, therefore, be classified into (Kusek and Rist 2004):

- Sequential complementarity,
- Information complementarity, and
- Interactional complementarity

*Sequential complementarity:*

Sequential complementarity comes in from the fact that monitoring information can generate questions that evaluation will have to address or evaluation information may give rise to new areas or domains of monitoring to be initiated.

*Information complementarity:*

Information complementarity arises from the fact that both M&E can use the same data, but answer different questions based on different analyses. Evaluation usually includes analysis of monitoring data. However, these data may not be adequate to provide reliable analysis and explanations on performance. In such cases, evaluation activities may engage in additional data collection, usually primary data collection.

*Interactional complementarity:*

Interactional complementarity refers to the fact that decision-makers make use of both M&E information in tandem to make decisions.

## 4 Participatory monitoring and evaluation

### 4.1 Nature of participatory monitoring and evaluation

Recognition of the benefits of participatory monitoring and evaluation (PM&E) is inspired from the dissatisfaction with top-down oriented conventional M&E. The main characteristics of the conventional M&E approach are that it is focused on measurement, is oriented to the needs of donors and policymakers (rather than participants or local people) and strives for objectivity (Estrella and Gaventa 1998). Moreover, conventional M&E is based on the tradition of scientific investigation and attempts to produce information that is necessarily objective, value free and quantifiable. Since outsiders usually do evaluations with the principle of enhancing 'objectivity', participants who may be affected by the findings of an evaluation have little or no input in the process, either in determining the questions asked and type of information obtained or in reflecting and using evaluative information.

Estrella and Geventa (1998) summarized the major criticisms of conventional approaches to M&E as follows:

- Costly and ineffective in terms of measuring and assessing achievements;
- Failed to actively involve stakeholders (beneficiaries and others who may be affected by the M&E);
- Made evaluation a specialized field and activity which is conducted and controlled mostly by outsiders and removed from the ongoing planning and implementation of development intervention;
- Used primarily as means to control and manage programs and resources, alienating intended beneficiaries and others involved in planning and implementation of development interventions from taking part in appraisal of the intervention; and
- Emphasized quantitative measures and tends to ignore qualitative information. Qualitative information may help provide fuller understanding of processes, outcomes and impacts.

In response to the problems associated with the conventional top-down approaches to M&E, new approaches to conducting PM&E evolved. These approaches aim to make M&E more responsive and appropriate to people's needs and real life contexts. The term PM&E is used to refer to an approach which focuses on collaborative learning and problem solving through the generation and use of knowledge. In line with this, Alur et al. (2005) defined PM&E as 'keeping track of changes with the community stakeholders'. PM&E is, therefore, a process that leads to corrective action by involving stakeholders at different levels in a shared decision-making process. In particular, PM&E involves bringing people at the grassroots and other stakeholders to actively participate in all stages of M&E of an intervention.

PM&E has emerged over the past 30 years based on the use of participatory methods in research and development. The recognition of the importance of PM&E arose from the trend in many agencies towards transparency, performance-based accountability, RBM, and the requirement to demonstrate success. Hence, participation has become a buzz word in development intervention. The concept of participation has become critical in assessing the needs of target groups and in implementation of interventions both in government and non-governmental organizations. With increased emphasis on the importance of participation in implementation of development interventions, there has also come a growing recognition about the importance of participation in M&E of development intervention. Hence, emphasis shifted away from externally controlled data seeking programs towards the recognition of locally relevant processes for gathering, analysing and using information.

The main arguments for PM&E are:

- Enhanced participation, especially of beneficiaries, in M&E helps improve understanding of the development process itself;
- Increased authenticity of M&E findings that are locally relevant;
- Improved the sustainability of the intervention, by identifying strengths and weaknesses for better management and decision-making;
- Increased local level capacity in M&E, which in turn contributes to self-reliance in overall implementation; and
- Shared experience through systematic documentation and analysis based on broad-based participation.

The major differences between conventional and PM&E are summarized in Table 1.

**Table 1.** *Difference between conventional and participatory evaluation*

	Conventional	Participatory
Who	External and internal experts	Community members, project staff, facilitator
What	Predetermined indicators of success, principally cost and production outputs	Joint identification of indicators of success, which may include production outputs
How	Focus on 'scientific objectivity'; distancing of evaluators from other participants; uniform, complex procedures; delayed, limited access to results	Self-evaluation; simple methods adapted to local culture; open, immediate sharing of results through local involvement in evaluation processes
When	Usually upon completion of project/program; sometimes also mid-term	More frequent, small-scale evaluations
Why	Accountability, usually summative, to determine if funding continues or not	To empower local people to initiate, control and take corrective action

Source: Adapted from Narayan-Parker (1993).

## 4.2 Uses of PM&E

PM&E can be used for different purposes, including planning and management of development interventions, organizational learning, impact assessment, and understanding stakeholder perspectives. PM&E can be used to achieve understanding of an ongoing development intervention in order to improve its planning and implementation. As a management tool, PM&E can be used to reflect systematically on a development intervention, and plan for future improvements. PM&E is also used to create learning processes to strengthen organizational learning. PM&E can help to assess organizational capacities and improve upon future implementation capacities. As such, PM&E helps people keep track of their progress, identify and solve problems by themselves and build on and expand areas of activity where success is achieved.

PM&E can also be used to assess the impact of a given development intervention. Assessing impacts of development interventions can help inform whether interventions are achieving their intended goals/outcomes, whether intervention objectives remain relevant over time, and whether or not the best strategies have been used. PM&E can also be used as a process that allows different stakeholders to present their needs, preferences, and expectations. Solicitation of such interests, needs and expectations would help planners of development interventions to incorporate beneficiary preferences into the planning of development interventions.



### 4.3 Fundamental principles of PM&E

Four key principles characterize PM&E: participatory, learning, flexibility and negotiation. It is important to ensure that M&E work adheres to these principles if it is to be truly participatory.

#### *Participation*

The core feature of PM&E is its emphasis on participation. PM&E acknowledges that there are multiple stakeholders who should participate in the M&E process. These stakeholders may include beneficiaries, project/program staff, government agencies and donors.

The issue of *who* initiated and conducted the M&E is one factor which characterizes the level and degree of participation in PM&E. Depending on who initiated, one can have M&E that is externally led, internally led or jointly led. In externally led PM&E, the efforts are generally initiated and organized externally and conducted by individuals or groups considered as having no direct involvement in the intervention. As the name indicates, in internally led PM&E, the efforts are carried out mainly by those directly involved in implementation. These include local people and staff members who are considered insiders. The joint PM&E combine approaches of internal and external M&E, and tries to assess from the viewpoints of both insiders and outsiders. The underlying objective in joint PM&E is to achieve a more holistic perspective and involve a more diverse set of stakeholders. The other point which characterizes participation in PM&E is the issue of *whose* perspective is particularly emphasized. This distinguishes the type of stakeholders who are emphasized by the M&E. In PM&E all major stakeholders, beneficiaries, or marginalized people can be emphasized.

#### *Learning*

Learning is the other distinguishing feature of PM&E. PM&E can be characterized as a process of individual and collective learning, through which people become more aware of their strengths and weaknesses. PM&E also enhances organizational learning. The learning process, it is hoped, would create conditions conducive to change and action. Learning in the context of PM&E emphasizes on practical or action-oriented learning. PM&E is a learning platform for those involved in the process. The process of learning is also a means for local capacity building, as participants gain skills which strengthen local capacity for planning, problem solving and decision-making. The process also gives participants of PM&E greater understanding of the various factors that affect the conditions and dynamics of the intervention, the basis for the success and failures of the intervention, and potential solutions or alternative actions. Participants of PM&E learn from their experience and gain the abilities to evaluate their own needs, analyse their own priorities and objectives, and undertake action-oriented planning. Overall, PM&E should serve to increase the analytical capacities of community members, and empower them to question, and become pro-active in development initiatives (Alur et al. 2005). PM&E can be undertaken monthly, quarterly and annually by different stakeholders at different levels, in order to achieve feedback, review and adjust the implementation of interventions.

#### *Flexibility*

Flexibility and experimentation are also regarded as integral parts of PM&E. PM&E should be contextual, which takes into account the local socio-cultural, economic, political and institutional contexts. There is no blueprint to carry out PM&E. The process of PM&E should be continually evolving and adapting according to intervention-specific circumstances and needs. This is because participating stakeholders can have new ideas on how to approach the M&E exercise. For instance, when participating stakeholders

come to the field, they may find that they have to adjust their plans because more people have come than expected, or for other reasons. Therefore, it is best to conduct participatory exercises by being flexible, while keeping in mind information required for effective M&E.

### *Negotiation*

PM&E is also perceived as a social process for negotiation between people's different needs, expectations and views. When multiple stakeholders are involved in the M&E process, negotiation is perceived as contributing towards the building of trust and changing perceptions, behaviours and attitudes among stakeholders, which affects the way they contribute to the intervention. PM&E can enable stakeholders to express their views more directly and more constructively.

## 4.4 Implementing PM&E

In order for the practice of PM&E be shared, replicated, compared and improved, it is important to systematize its practice. Although there may be variations in the way PM&E is practised, it is important to develop common guidelines and methods that characterize PM&E. Generally agreed answers are needed to the following questions: What are the key steps in PM&E process? What tools and techniques should be used? Who should be involved and how?

The major steps that should be followed to practice PM&E are not that much different from the ones in conventional M&E. However, the approaches used in each step varies from the conventional top-down approach to M&E, in that PM&E provides space for real stakeholder participation. This section briefly describes the major steps and highlights the issues that need to be considered to make the M&E participatory.

The major steps in PM&E are:

1. Establishing a PM&E team
2. Planning the PM&E process
3. Setting up of objectives and indicators
4. Developing data collection methods
5. Collecting data
6. Analysing data
7. Report writing
8. Dissemination of PM&E findings

While establishing a PM&E team, it is important to ensure that the team composition reflects the diversity of stakeholders. The planning stage is critical to the success of PM&E. The planning process may require negotiation and joint decision-making. At the planning stage, space must be allowed for different stakeholder groups to get together and air concerns, needs and expectations. At this stage, questions such as what information should be collected? For whom? Why? And how the information would be incorporated into planning and implementation of an intervention? should be answered.

Identifying the objectives and indicators can at times be the most difficult part in the PM&E process. Consultative process among the various stakeholders to determine objectives and indicators is considered by many as critical to carrying out successful PM&E. Usually, much more attention is given to indicator development, and less to determining the objectives of PM&E. However, it is important to first determine the objectives of the process before proceeding to indicator development. In order to determine the objectives of PM&E, one must know who will be the end-users of the information, and

how the results are to be used. End users can be direct beneficiary community members, staff of the organization implementing the intervention, donors, development agencies, research organizations, policymakers at different levels, or even indirect beneficiaries in a community. What is to be monitored and evaluated and how will be based on who needs the PM&E results and information. Who needs the information should also be determined based on a consultative process among stakeholders.

After objectives are established, then indicators can be developed (see section six for more information about indicators). The indicators guide the type of information to be collected. Once information needs and objectives are identified, data collection needs to be planned and implemented. There are a wide range of tools and techniques that can be used for data collection (see section eight for details about data collection and analysis). The issue of participation is also equally important in the next steps of analysing and assessing the finding, reporting and using the findings. Stakeholder groups should engage in critical reflection and thinking about the problems and constraints, the success and outcomes of their efforts and activities which they have undertaken. Data analysis can also be done in a variety of ways. There are basically no major differences in report writing and information dissemination between the conventional M&E and PM&E (see section nine).

In general, if the system of M&E is to be truly participatory, the first step is to make sure that the issues to be monitored and the indicators that will be used are identified, discussed and agreed by the relevant stakeholders. At each stage of PM&E, deciding who should be involved and on what terms is the most critical activity for sustaining the PM&E process. Relevant stakeholders should also be involved in deciding how often progress should be reviewed, who should do it, using what resources and method etc.

## 5 Selecting results (outcomes) to monitor and evaluate

In order to successfully measure performance, result expectations should be clearly defined for different level of the results chain. The information needs of stakeholders from M&E can be diverse. Some stakeholders need operational information related to the activities and outputs of the organization while others may need information on results (immediate outcomes, intermediate outcomes and ultimate outcomes). In view of these, careful consideration should be taken when deciding what to be tracked, documented and analysed using the M&E system. Therefore, the first step in building an M&E system is the identification of operational issues and expected results that need to be monitored and evaluated. The basis for identifying the information needs should be the results chain. Each level of the RC has unique information needs. Therefore, it is necessary to start by identifying the information needs in relation to each level of the RC. Information needs to measure inputs, activities and outputs is straightforward. However, careful thought is needed to determine how to measure results.

The information needs at the immediate outcome level is related to changes in awareness/skill and access to services and inputs. The information needs about intermediate outcome relates to whether, how and to what extent outputs have been used by intended beneficiaries. At the ultimate outcome level, the information needs relate to how the lives of beneficiaries have changed as a result of the intervention, what effects have there been on the environment and what changes have occurred on the social aspect of a community (i.e. gender, HIV/AIDS, climate change etc.). At ultimate outcome level, information may be linked to sectoral goals, national development plans, or international plans (e.g. MDG). For example, the ultimate outcome statement of the IPMS project which reads as ‘improved agricultural productivity and production within functional market-oriented agricultural production systems in and beyond the pilot learning districts’ is linked to the government agricultural development goal.

### 5.1 Qualities of good result statements

A result statement should be SMART: specific (specify the nature of the change, the target groups, and the target region), measurable (measured by using indicators), achievable (realistic), relevant (answers identified need), and time bound (time lines by which results are to be achieved are specified). CIDA (2009) gave further clarification about qualities of good result statements (see Box 1).

### 5.2 Process of selecting result statements

Steps in choosing outcomes to monitor and evaluate include identification and involvement of key stakeholders, identification of major concerns of stakeholder groups, translating problems into statements of possible outcome improvements, ensuring that outcome statements are disaggregate, and developing a plan of how the intervention will achieve the outcomes (Kusek and Rist 2004).

#### Identifying and involving key stakeholders

Since the information needs of different stakeholders of an intervention can be different, it is important first to identify the key stakeholders before identifying outcomes. Questions such as Who are the key stakeholders involved around an issue? How are they characterized? Whose interests and views are to be given priority? should be answered. After the ‘right people’ are identified they have to be

involved in selecting the results statements. Though the 'right people' may vary based on the type of project/program/policy at hand, it may include experts from the relevant fields, beneficiaries, partner organizations, donors etc.

### Box 1: Characteristics of quality result statements

- Is the statement simply worded? The result statement should be clearly stated and easy to understand.
- Does the result statement contain only one idea? If it contains more than one idea, think of splitting into separate statements.
- Was the result statement drafted in an inclusive, participatory fashion? In general, the process of selecting information to be monitored should be participatory involving key stakeholders. Note that the information needs of various stakeholders from the M&E system varies. For example, in a project which tries to improve rural livelihood by introducing new agricultural technology, donors may be more interested in the number of women who benefited from the intervention, government officials may want information about the contribution of such technology for poverty reduction, beneficiaries may be interested in the net income generated from the technology and researchers may be more interested in the change of productivity by using the new technology. Therefore, it is important to make sure that all the voices are heard and that the expected outcomes are shared with all involved is essential.
- Are the results truly gender sensitive? Do they address the concerns, priorities and needs of women and men, girls and boys?
- Do the results consider environmental implications? Will results be sustainable?
- Does the result statement include an adjective and tells:
  - What? Does the result statement describe the type of change expected using an adjective that is drawn from a verb and that indicates direction (increased, improved, strengthened, reduced, enhanced)?
  - Who? Does the result statement specify the target population or beneficiary of the intervention? Does it specify the unit of change (individual, organization, and group)?
  - Where? Is the location or site where the result will occur specified?
- Can the result be measured? Can the result be measured by either quantitative or qualitative performance indicators? Can performance indicators that will measure the result be easily found, collected and analysed?
- Is the result realistic and achievable? Is the result within the scope of the project/program's control or sphere of influence? Is there an adequate balance between the time, resources allocated and the expected reach and depth of change expected? Are the results at the immediate and intermediate level achievable within the funding levels and time period of the project/program? Is the result (immediate and intermediate outcome level) achievable during the lifecycle of the investment?
- Is the result relevant? Does the result reflect the needs of the target group and will it support higher level developmental change in the strategies or programs it supports?

Source: CIDA (2009).

## Identifying major concerns of stakeholder groups

Participatory M&E requires that concerns of relevant stakeholders be incorporated in the M&E process. Hence, the interests of various stakeholder groups must be solicited using quick information gathering techniques, such as brain storming sessions, focus group interviews, key informant interviews etc. A major problem which would arise in doing this is that the number of issues to be monitored could increase to the level which might be difficult to handle with the available resource for M&E. In this regard, continuous negotiation and discussion have to be made to reach a consensus and to limit the number of issues to be monitored to the most important ones.

## Translating problems into statements of possible outcome improvements

Selecting ultimate outcomes is done by identifying the problem a project/program/policy intends to address. An outcome statement enables one to identify the path and destination of an intervention. It is preferable that outcomes be prepared in positive statements than in negative statements, as it may be easier to rally stakeholders behind positive statements.

## Ensuring that outcome statements are disaggregated enough

It is important that outcomes be disaggregated sufficiently to capture only one improvement area in one outcome statement. Operationalizing M&E is easier if outcome statements are disaggregated to the lowest level possible. Generalized outcome statements make it difficult to develop indicators, baselines and targets. Answering questions such as for whom, where, how much and by when helps to disaggregate outcome statements. It is only disaggregated outcomes will help to know if the results are achieved or not. Disaggregating or scaling down outcomes into subcomponents helps identify specific measurable indicators.

## Developing a plan of how the intervention or organization will achieve outcomes

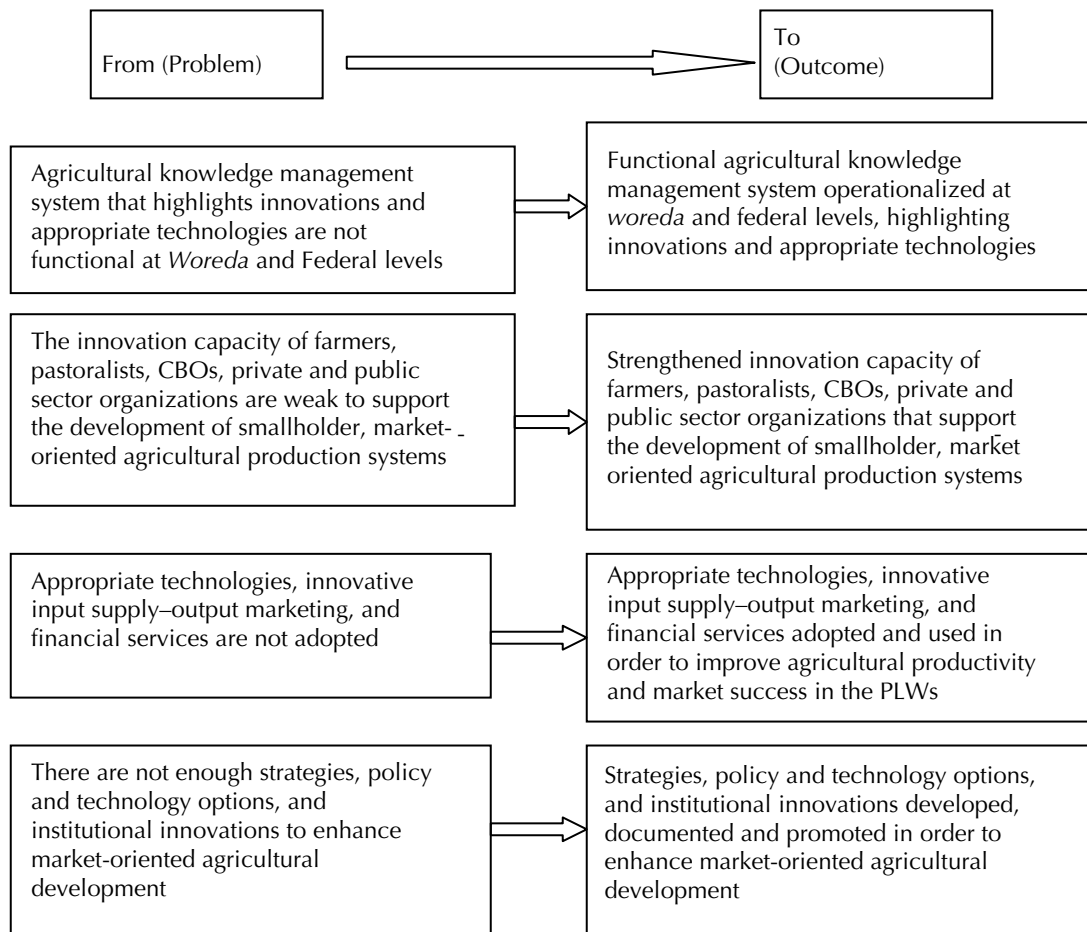
Gathering information about inputs, activities and outputs is straightforward, since the logical connection from inputs to activities, to outputs is very direct. However, it must be noted that completing all the activities and achieving all the outputs is not the same thing as achieving the desired outcomes. This means that concerns and problems should be recast into solutions. Figure 7 illustrates the translation of problems into solution statements using the IPMS project as an example.

In general, certain key questions need to be answered in choosing what to monitor and evaluate (IFAD 2002).

1. Ensure that information will be collected for each level of the results chain (outcomes, outputs, activities and inputs)
2. Consider the information needs of different stakeholders in a participatory way
3. Include information that helps to answer the five core evaluative questions:
  - a. Relevance: whether the intervention is still necessary in relation to the target group priorities?
  - b. Effectiveness: whether the planned outcomes, outputs and activities are being achieved?
  - c. Efficiency: whether inputs (resources and time) are used in the best possible way to achieve outcomes?
  - d. Ultimate outcome (impact): to what extent the intervention has contributed towards its longer term goals
  - e. Sustainability: whether there will be continued positive impacts as a result of the intervention once it has been finished

Note that implementation focused M&E systems focus on the issues of effectiveness and efficiency; whereas RBM&E systems give more emphasis to the issues of relevance, impact and sustainability.

4. In addition to information that helps to check targets, include information that can help to explain progress. Knowing why (or why not) something happened gives a clearer base for decision-making and interpretation of the result.
5. Look out for positive and negative unintended impacts.
6. Stick to the principle of 'less is more.' Information should be collected only if it is relevant and will be used.



Source: IPMS project performance framework (2005).

**Figure 7.** Outcome statements derived from identified problems or issues, IPMS project.

## 6 Selecting key performance indicators

After the operational issues and the result expectations to be monitored are selected in participatory manner, the next step in the process is selecting indicators. In order to measure the accomplishment of activities and achievement of outputs and outcomes, first their measurable indicators have to be identified. It is through these key indicators that one measures the change and determines whether the intervention is on the right track or not.

A performance indicator is a quantitative or qualitative unit of measurement that specifies what is to be measured along a scale or dimension (CIDA 2009). A performance indicator is neutral; it does not indicate a direction or change nor incorporate a target. Outcome indicators are not the same as outcomes. Indicators are the quantitative or qualitative variables that provide a simple and reliable means to measure achievement of results, to reflect the changes connected to an intervention, or to help assess the performance of an organization against the stated result (Kusek and Rist 2004). In simple words, indicators are clues, signs or markers that measure one aspect of an intervention and show how close the intervention is to its desired path and outcomes. That is, they represent a simplification or approximation of the issues to be monitored and help to measure and communicate changes.

Indicators should be developed for all levels of the M&E system. In other words, indicators should be developed to monitor and evaluate progress and performance with respect to input use, accomplishments of tasks and activities, achievement of outputs and outcomes. Indicator development is a core activity in the M&E process and must be done with care, since indicators determine the type of M&E data that will be collected. Similar to the process of identifying outcomes, the views and ideas of relevant stakeholders should be taken into account when selecting indicators. Attempts to accommodate interests of various stakeholders can increase the number of indicators beyond the budgetary limits of M&E system. However, a set of minimum indicators that measure the outcome directly should be included.

For example, in the IPMS project one of the outcomes expected to be realized is ‘functional agricultural knowledge management system operationalized at district and federal levels, highlighting innovations and appropriate technologies’. But how do we know whether and when we achieve this outcome? In order to measure this outcome one indicator used by the project is ‘frequency of information exchange among stakeholders’. Therefore, by counting the number of information exchanges among stakeholders, it is possible to measure the realization of ‘functional knowledge management system’.

Setting indicators help to easily provide feedback and identify success/achievement, concern that need to be improved, and to check whether the intervention is on the right track. In addition to the use of indicators for monitoring the status of activities, outputs and outcomes, indicators can give us a wealth of performance information about the process of and progress towards achieving results. Managers can use the information from indicators to identify deviance from plan in achieving various levels of results.

Selecting simple and reliable indicators for results is more difficult than selecting indicators for activities and outputs. This is because changes at results levels are usually a product of changes in different component of the intervention. For example, in the IPMS project, the ultimate outcome level change is ‘improved agricultural productivity and production within functional market-oriented agricultural production systems’. The project has four objectives: knowledge management, capacity building, commodity development, and research. Each one of these objectives is expected to make its own



contributions to the achievement of the highest level change (ultimate outcome). In such cases using a single indicator or few may not be adequate to understand the changes.

## 6.1 Characteristics of good performance indicators

Good indicators possess certain characteristics. These characteristics can be described by the acronym CREAM (Kusek and Rist 2004). CREAM stands for clear, relevant, economical, adequate and monitorable. Below we give brief description of each of these characteristics.

### *Clearness:*

Performance indicators should be clear, direct and unambiguous. Clear indicators are more precise and coherent which help to have a better focused measurement strategies. The more precise one can make each indicator, the less likely it is misunderstood by the people involved in data collection and analysis. One can make an indicator precise by indicating (1) specific target group to which the indicator will be applied (2) specific unit(s) of measurement to be used for the indicator (3) specific timeframe over which it will be monitored (4) defined qualities for the adjectives used and specific location in which indicator will be applied. For instance, suppose training is given to experts working in the Offices of Agriculture and Rural Development (OoARD) on broad bed maker (BBM) so that they can train farmers. Here one can use the indicator 'perception of training participants'. But this indicator is not clear; we can make the indicator clearer by restating it as 'perception of experts who attended training on BBM about how the training helped them to train farmers'.

### *Relevance:*

Indicators should also be relevant (appropriate) to the desired result expectation and they should not be affected by other issues which are indirect to the result. Relevance of indicators relates to the usability of the indicator to measure progress.

### *Economical:*

One should also think about the economic cost of collecting data on that indicator. Feasibility of an indicator is related to accessibility of data. If data on an indicator can be obtained at reasonable cost and effort then that indicator is feasible and economical. When we talk about feasibility, both financial and technical feasibility should be considered. Financially, one should use the budget limit to decide what one 'needs to know'. In addition, technically one has to confirm the availability of human capacity to assess the indicators. In this case, it is important to ensure that there is access to resource persons with the skills to carry out the necessary analysis. Choose performance indicators that provide the best possible measurement within the budget available and wherever possible, use existing sources and data collection methods.

### *Adequacy:*

Indicators ought to be adequate, i.e. they should not be too indirect, too much of a proxy, or so abstract that assessing performance become complicated and problematic. Adequacy also refers to representativeness of the indicator. An indicator is fully representative if it covers the most important aspect of the objective that is being tracked. At higher level of objectives, several indicators should be selected to ensure the representativeness of the indicators. If an indicator is not adequately representative, it is always good to consider additional indicators.

*Monitorable:*

Monitorability of indicators refers to whether an indicator can be independently validated or verified. Indicators should also be reliable and valid to ensure that what is being measured at one time is what it can be measured at later time and that what is measured is actually what is intended to measure. If any one of these five criteria is not met, formal performance indicators will suffer and will be less useful.

In addition to the CREAM criteria, some authors suggested different sets of criteria. Roche (1999) claimed that when indicators are used more as specific examples of change, different characteristics become important. In this context, he outlined another set of characteristics for indicators using the acronym SPICED (subjective, participatory, interpreted, cross-checked, empowering, and diverse). Below we give brief description of these criteria.

#### *1. Subjective*

Informants have a special position or experience that gives them unique insights which may yield a very high return on the investigators time. In this sense, what may be seen by others as 'anecdotal' becomes critical data because of the source value.

#### *2. Participatory*

Indicators should be developed together with those best placed to assess them. This involves beneficiaries, local staff and other relevant stakeholders.

#### *3. Interpreted*

Locally defined indicators may not mean much to other stakeholders, so they often need to be explained.

#### *4. Cross-checked*

The validity of assessment needs to be cross-checked, by comparing different indicators and progress, and by using different informants, methods, and researchers. This characteristic is more critical when the bias of sources is polarized.

#### *5. Empowering*

The process of setting and assessing indicators should be empowering in itself and allow groups and individuals to reflect critically on their changing situation.

#### *6. Diverse*

There should be a deliberate effort to seek out different indicators from a range of groups, especially men and women. This information needs to be recorded in such a way that these differences can be assessed over time. This characteristic is more important when the intervention has differential effect on gender, age, ethnicity etc.

## 6.2 Types of indicators

Indicators are classified into quantitative, qualitative, proxy indicators and pre-designed indicators. Below we give brief description of each of these types of indicators.

### *Quantitative vs. qualitative indicators*

In establishing M&E system, it is recommended to start with a simple and quantitatively measurable system rather than inserting qualitatively measured indicators upfront. Quantitative indicators are discrete measures or indicators that are expressed numerically (number, mean, median, percentile, and ratio). Examples of such indicators are number of farmers who adopt a technology, ratio of women-to-men in extension service, percent of farmers who adopted a given technology, percent of farmers above a certain income level, percent of farmers who contribute to collective action etc. Outcome indicators are often expressed as the number or percent of something. However, it should be stressed that using both numbers and percentages for a given indicator provides more complete information than just using number or percentage. The quantitative indicators directly measure the status or change of specific variables. For example, crop yield in kg, kilometre of irrigation canal constructed are quantitative indicators which provide direct numerical results.

Qualitative indicators, on the other hand, are variables that measure an issue based on qualitative assessment. Qualitative indicators are measures of an individual or group's judgement and/or perception of congruence with established standards, the presence or absence of specific conditions, the quality of something or the opinion about something (CIDA 2009). Many of the qualitative indicators use adjectives such as successful, adequate, equitable, good, effective, participatory, empowered and well functioning. Qualitative indicators can also use scales or ranks, such as highly, moderately, poorly, adequately etc.

Qualitative indicators are collected by asking people to express their opinion, judgement, perception or explain what they have observed. They are reported using qualitative assessments such as congruence with, quality of, extent of and level of. Qualitative indicators provide insights into changes in institutional processes as well as attitudes, beliefs, motives and behaviours of individuals. A qualitative indicator might measure perception, such as the level of farmers' satisfaction with regard to a new extension approach. Qualitative indicators might also include a description of behaviour, such as the level of mastery of a newly learned skill. The advantage of qualitative indicators is that they can capture things such as perception and quality which are difficult to express in numbers (quantitatively). However, data collection, measurement and analysis for qualitative indicators is more time consuming, and likely to be less accurate and prone to subjectivity.

For the sake of simplification, information from qualitative indicators can be quantified. For example, opinion of people can be categorized and counted to express it numerically. But making quantitative information qualitative is not possible as it is difficult to extract opinion from a number. In general, there should be a balance about the number of indicators between qualitative and quantitative. Qualitative indicators provide more in-depth information. These two types of indicators can be used in complementary way. For example, we can take opinion of target group on something by listing their opinion, but also we can use a more quantitative approach and ask the target group to report on the issue quantitatively.

### *Proxy indicators*

It may not always be possible to develop precise and direct indicators. Some outcome indicators may be difficult to measure directly. In such instances, one can strive to use approximate measures. However, use of indirect or proxy measures should be limited only to situations when data for direct indicators are not available, when data collection may be too costly or beyond the available budget, when data cannot be collected at desired intervals, or when data collection is not feasible at all. In

using proxy indicators, it is important to ensure that the proxy is giving at least approximate evidence on the performance.

### *Pre-designed indicators*

When indicators are constructed independently of the context of an individual country, sector, organization, project/program/policy, the indicators are known as pre-designed indicators. A number of international development organizations may construct indicators to track development changes. Examples of pre-designed indicators include the Millennium Development Goals (MDGs), the United Nations Development Program's (UNDP's) Sustainable Human Development Goals, The World Bank's Rural Development Indicator, and the International Monetary Fund's (IMF's) Financial Soundness Indicators.

MDGs contain eight goals, with corresponding targets and indicators assigned to each. UNDP established the Human Development Index (HDI) in 1990 to measure the quality of life in all countries. The World Bank's Rural Development indicator contains indicators for rural wellbeing, improvements in the rural economy, development of rural markets, and others. The IMF uses indicators of financial soundness.

The advantages of using pre-designed indicators include possibility to aggregate across similar projects/programs/policies and possibility to make cross-country comparisons. The disadvantages of using pre-designed indicators include difficulty to address country specific goals, top-down nature and lack of stakeholder participation. It is important to note that pre-designed indicators may not be relevant to a given country or organizational context.

## 6.3 Constructing indicators

Indicators determine the type of information to be collected for M&E system. Indicator construction needs time and should be done with care and in a participatory manner. It should be done by competent personnel who have expertise in the issues covered by the M&E system. Indicators have to be constructed to meet specific needs. Since the objective is to achieve measurement of the results chain, indicators should reflect each element of the results chain. Although indicators should be identified across the performance framework, from resources through ultimate outcome, it should be noted that RBM emphasizes measuring achievement of results more than the use of resources (CIDA 1999).

Indicator selection guides the subsequent operation by indicating what information/data to look for. If we do not assess the issue with the right measurement, we may not be able to get the right information needed. Therefore, it is good to ask the right question for each of the issues selected for M&E.

Stakeholder participation in selecting indicators helps the incorporation of stakeholders' interests and concerns, and improves the usability of M&E information. Therefore, it is important to distil stakeholders' interests into good, usable indicators. Thus, the different level of results should be disaggregated to make sure that indicators are relevant across the concerns of different stakeholder groups and not just for a single stakeholder group. It is important to be thoughtful about the number of indicators chosen, because each indicator will need data collection, analysis, and reporting system behind it. Monitoring too many indicators, given the limited capacity to do so, is difficult and decreases the quality of the M&E information. Moreover, too many indicators can also negatively affect the response rate. It also gets in the way of the 'real' work of implementation and complicates things. Therefore, it is very important to reduce the number of indicators to the minimum necessary that meet key management,

learning, reporting and decision-making needs. In addition to reducing the number of indicators, the frequency, and level of detail can also be reduced to ensure manageable size of indicators. In general, it is good to include only the information required to improve decision-making. This calls for the regular revisiting of the indicators selected for monitoring.

Although indicators can be adopted or even dropped, dropping or modifying indicators should be done with caution and consultations. When indicators are changing, baselines against which to measure change or progress also needs to be changed. In order to minimize revision, indicators should be thought through carefully when they are first established.

## 7 Setting baselines and targets

After formulating result expectations and selecting key performance indicators, the next step is setting of baselines and targets.

### 7.1 Setting baselines

Progress cannot be assessed meaningfully without baseline figures in place. In this regard, an initial basis for comparison is important to assess the changes over time and to ascertain if these changes are the result of the intervention concerned. Therefore, it is important to have information about the initial starting situation before any intervention has taken place. This kind of information is known as baseline information.

A baseline is qualitative or quantitative information that provides data at the beginning of, or just prior to, the implementation of an intervention. It is the set of conditions existing at the outset of a project/program/policy. Baseline information establishes where we are at present relative to the outcomes we are trying to achieve (Kusek and Rist 2004). In other words, the baseline is the first measurement of an indicator. It provides a benchmark against which future performances can be tracked. Baseline information also informs decision-makers about current situations and patterns of performance before embarking on projecting targets for a given project/program/policy.

Establishing baselines involves collecting data on indicators to show the initial position of a situation before the intended intervention. Baseline can be used to learn about current or recent levels and patterns of performance. For example, one can take an average of the last three years' performance or the first year's performance or the average trend over the past six months as a baseline. The important thing is to decide which option is feasible and gives clear view of the baseline situation.

In establishing a baseline, it is important to collect only data that is intended to be used by decision-makers and other relevant stakeholders. Information that will not be used should not be collected. Baseline information should be collected for the result expectations and their selected indicators discussed in the previous sections. Therefore, collect baseline data that relate directly to the indicators that are identified.

This baseline information serves as a point of comparison. In order to do comparison we can have two options. One option is comparing the situation 'before the intervention' with the situation 'after the intervention'. However, such comparison fails to account for the changes that would occur without the intervention, thus may lead to erroneous conclusions. The second option is to compare changes with and without an intervention. The with-and-without approach compares changes in the intervention area with similar locations outside the intervention area. Such comparison captures the incremental changes that occur as a result of the intervention.

### 7.2 Setting targets

After gathering baseline data on indicators, the next step is to establish targets for the different levels of results. A target is '... a specified objective that indicates the number, timing and location of that which is to be realized' (IFAD 2002). In essence, targets are qualified or quantified levels of the indicators that an organization or a project/program wants to achieve by a given time. In other words, a target specifies a particular value for a performance indicator. Result targets establish what can be achieved in a specific time frame toward reaching the expected results. For example, one target might be '50% of rural households in the intervention areas should be able to sell 75% of their products by the year 2009'.

Targets are established for each indicator by starting from the baseline level, and by including the desired level of improvement in that indicator. In doing so, one should not be too ambitious and it is important to be realistic about the results that are feasible to achieve given the contextual constraints and past experience in a particular sector. Targets may be either quantitative or qualitative, depending on the nature of their indicators. While targets for quantitative indicators will be numerical, targets for qualitative indicators will be descriptive. Quantitative indicators identify how much of a change is expected from year to year. Indicators that focus on changes which are not easy to describe in quantitative terms can also be selected. Targets should be set at the design/planning stage but should be periodically reviewed based on actual performance.

There are a few critical information sources that need to be consulted in establishing targets (Box 2). Such information relate to baselines, historical trends, stakeholder expectations, expert opinions and research findings, and accomplishment of similar projects/programs (USAID 1996).

### Box 2: Establishing targets

- Baseline data indicating the situation at the beginning of an intervention. When such data is not available, management should include an activity to collect it from the start.
- Historical trends in the indicator value over time. What pattern of change has been evident in the past? Is this pattern likely to continue?
- Stakeholders' expectations of progress. Exploring the achievement expectations of local counterparts such as project implementers and managers may be useful to provide a realistic idea of what can be achieved.
- Expert judgements and research findings. Experts knowledgeable about the sector and local conditions as well as research findings are other useful sources of information for target setting.
- Accomplishments of similar programs. Information on what is being done in the program sector under similar conditions by other agencies and organizations that have a reputation for high performance is excellent input to the target setting process.

Source: USAID (1996).

In addition to ensuring critical information sources are consulted, several important considerations should also be made before performance targets are set. First, the available resources over a specific time period to arrive at the performance target should be the basis to formulate targets. Therefore, it is important to know the starting point (the baseline) and the available resources to achieve results before deciding on targets. Organizational capacity, expected funding and resource levels (budgets, personnel, funding, facilities) including internal as well as external funding sources need to be projected for the intervention period.

Second, previous performances should be considered in projecting new performance targets. One might observe how an organization or a project/program has performed over the years before projecting future performance targets. Past performances could give a reasonable idea of how an organization might perform in the future. For this, the baseline information should be consulted.

Third, since it is difficult to foresee the future precisely, and forecast risk factors accurately, it is wiser to limit setting targets to periods of less than five years. If the intervention would last for longer than five years, it may be better to roll targets over the intervention period of five years or less. In most cases targets are set annually, but can also be set quarterly or they can even be set for longer period.

## 8 Data collection and analysis

Once decision is made on information/data needs and the associated indicators to be used, data collection for M&E should proceed. In M&E system, considerations about data collection should be made starting from the process of indicator selection. As part of indicator selection one should consider source of data, data collection methods, frequency of data collection and who will be responsible for data collection, analysis and reporting. Such early considerations would help the indicator selection to be based on the organizational reality and may point out issues of the data system that should be improved in terms of data collection strategy and staff capacity.

Only data that is going to be used for making decisions should be collected. Since performance information should serve as management tool, it may be a waste of resources to collect information that decision-makers may not use. Data quality may also be compromised if we intend to collect data that is not directly relevant to decision-making, and data collection cost and time may escalate.

In designing data collection for M&E, it is important to evaluate the data collection methods in terms of data accuracy, reliability, validity and timeliness. Accuracy refers to the level of precision with which the data would be collected. Reliability refers to the extent to which a data collection system is stable and consistent across time and space. In other words, reliability of a data collection system implies that measurement of the indicators gives the same meaning every time the measurement is repeated. Validity refers to the indicators' direct measurement capacity of what is intended to be measured. In other words, validity refers to how well the data collection instrument corresponds to what is intended to be measured. It is important to be clear for each indicator regarding what sources of information can potentially supply the relevant data. Timeliness is about accessing the needed information in time to make timely decision-making possible. If the data collected through the M&E system is not made available to decision-makers when they need it, the information may become obsolete.

To meet these criteria, good M&E systems should include a clear data collection and analysis plan which details the type of data, data collection methods and instruments, frequency of data collection, methods of data analysis as well as reporting and dissemination procedures. Time spent on planning the data collection may have high payoff in saving time and data quality and completeness. It is important to resist the temptation to embark on data collection before making adequate planning.

### 8.1 Source of data and frequency of data collection

#### 8.1.1 Sources of data

Data can be collected from different sources. Data sources may be individuals, groups, organizations or documents from which data about indicators will be obtained. Most of the time, target populations provide the bulk of data for M&E. Partner organizations (governmental, non-governmental and CBOs) as well as members of the private sector can serve as a source of information for M&E. Experts who are knowledgeable in the topic can provide useful information. Moreover, individuals who have experience in the aspects of the development intervention can serve as key informants. Researchers and previous studies conducted on aspects of the intervention can supply raw data or processed information.

Data for indicators can be categorized as primary and secondary. While primary data are collected directly by the M&E system for M&E purpose, secondary data are those collected by other organizations for purposes different from the M&E. Secondary data may be more cost effective, or it may be the only



possibility when primary data collection is not practical. However, caution should be taken in using secondary data. The validity and reliability of secondary data must be checked carefully. If the secondary data was collected through surveys, the sampling procedure and sampling unit, and frequency of data collection must be carefully examined to ensure that they are consistent with the M&E data needs. Since the source of performance data is important to the credibility of reported results, it is important to incorporate data from a variety of sources to validate findings. In general, in choosing the data sources, access, quality as well as feasibility and cost effectiveness should be considered.

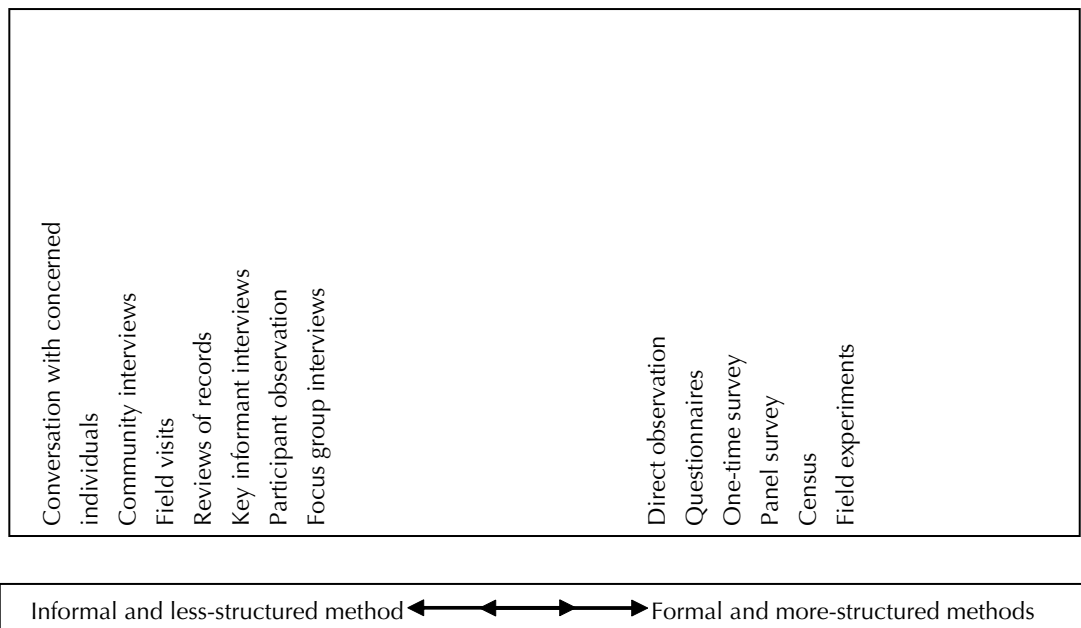
### 8.1.2 Frequency of data collection

Data can be collected at different time intervals. Frequency of data collection refers to the timing of data collection; how often will data/information about each indicator be collected and/or validated? Will information/data be collected regularly (e.g. quarterly or annually) as part of ongoing performance management and reporting, or periodically, for baseline, midterm or final evaluations? In general, frequent data collection means more data points. More data points enable managers to track trends and understand the intervention dynamics. The more often measurements are taken, the less guess work there will be regarding what happened between specific measurement intervals. But the more time that passes between measurements, the greater the chances that events and changes in the system might happen that may be missed. However, more frequent data collection also means more additional cost. Therefore, a balance should be made between cost and frequency.

The frequency of data collection usually varies across the results chain depending on time of occurrence of the expected changes. Activities are undertaken continuously and the changes in these variables are frequent. Similarly, outputs as the direct result of activities can be observed frequently. On the contrary, outcomes are expected to be seen in the short-, medium- and long-term. The frequency of data collection also follows this pattern. For example, in the IPMS project data on inputs and activities are collected and reported biannually. Output and outcome information are collected annually starting from the second and the third years of implementation, respectively. On the other hand, impact (ultimate outcome) information is collected at the end of the project year.

## 8.2 Data collection methods and instruments

Once the sources of data and the frequency of data collection are known, then the methods and instruments for data collection should be determined. A method is an established systematic way of carrying out a particular task. M&E system uses different methods or procedures to collect data. Data collection methods represent *how* data about indicators is collected. Methods of data collection for M&E system include discussion/conversation with concerned individuals, community/group interviews, field visits, review of records, key informant interviews, participant observation, focus group interviews, direct observation, questionnaire, one-time surveys, panel surveys, census and field experiments (Kusek and Rist 2004). These methods range from highly informal and less structured to highly formal and structured (Figure 8). The choice of data collection method depends on the type of indicator, the purpose of the information being gathered and the frequency of data collection.



Source: Kusek and Rist (2004).

**Figure 8.** *Methods used for data collection.*

## Conversation with stakeholders

M&E personnel may start from conversation with stakeholders in order to get a general idea of the intervention, the role of the stakeholders, their experiences, and perceptions about the performance of the intervention. Such conversation can be done on a one-on-one basis or with a group of stakeholders. This method is usually the very first step in the data collection process.

## Community interviews

Community interviews are conducted with a group of community representatives in the intervention area or part of the intervention area. Depending on the size of the intervention area, several community interviews can be conducted. The purpose of community interviews is to solicit community perspectives on the performance and results of the intervention in the area. Such interviews can be guided by checklists and open-ended questions. But, the interviews remain fairly open and flexible to accommodate emerging issues.

## Field visits

Field visit is one means to collect data. It is a visit made to an area, usually by a group of experts. The length of the visit may vary from one day to more than a week depending on the size, complexity of the site, and the availability of experts. Being already informed to some extent about the site from the review of previous studies and from the secondary data, the group member can form general impressions about the issue under investigation and attempt to gain more knowledge concerning special areas.

## Review of records

This method is used to understand the historical evolution and performance of a project/program or organization through its documentation in the form of written, electronic, photographic or video materials. It can also provide a good background to help explain whether changes are occurring. Such

an initial review of the literature can also help in identifying key issues that needs to be addressed in a further M&E analysis. The disadvantage of this method is that the limitation of the information by what documentations available and accessible, how it has been presented and by whom it is presented, and how it has been stored.

### Key informant interviews

Key informants are individuals who have substantial knowledge about the performance indicators either because they are beneficiaries of the intervention or major stakeholders in the intervention. Key informants can be farmers, experts, administrators, or development practitioners. The advantage of this method is that it is quick and less expensive; a carefully administered key informant survey gives reliable information.

### Participant observation

Participant observation is the process of systematically and continuously documenting community and individual behaviours, without disrupting the process, people or locations being observed (Friis-Hansen et al. 2000). It can be used to gain information about how patterns of behaviour and social interactions, livelihood strategies, and environmental conditions may have changed as a result of the intervention.

This method helps to obtain useful and timely information by observing what people do and to help make decisions on improving performance. Usually this method is used to complement information collected using other methods. This method is especially suitable for understanding processes which are difficult to grasp in an interview context. Moreover, it can be used to understand the context in which information is collected and help to explain results.

Observation can be made more effective by viewing it as a valid method and structuring its use. Much can be learned by watching what people actually do. Useful information and new insights can often be gained from such observation that would otherwise not be obtained. If done well, it can permit a deeper understanding of relationships within communities but also between a community and other organizations. This method is a common research method for social issues and processes. Direct observation is useful for validation in monitoring as it can be used to cross check responses received by other methods.

In participant observation, biases can be created due to the observer, the way the observer influences the behaviour of the observed or the observed situation, thus hampering the objectivity of the observer. These biases can never be eliminated entirely. Therefore, direct observation as a systematic M&E method should only complement other methods. Asking several people to undertake observations in the same manner can help confirm observations or identify differences and so increase the quality of the data.

### Focus group interviews/discussions

A focus group discussion is a discussion made by a panel of 8 to 12 respondents led by a moderator or facilitator. The moderator uses group dynamics principles to focus or guide the group in an exchange of ideas, feelings, and experiences on a clearly understood topic. Focus group interviews help to collect general information, clarify details or gather opinions about an issue from a small group of selected people who represent different viewpoints. It can also be used to build consensus. Analysis of gathered information attempts to discern patterns and trends that develop among participants, as well as across

focus groups. For M&E, focus groups are good for assessing opinions of change, assessing the quality of project/program services or service providers, and identifying areas of improvement.

## Direct observation

Direct observation techniques are the most reliable way to assess if your users are satisfied with the service projects/programs or organizations are providing, and to evaluate their actual performance and compliance with standards. Direct observation techniques evaluate user perceptions and satisfaction, and bring the user and the provider into close contact. Direct observation should be used with great sensitivity lest it create tension rather than support solutions. Participant observation is one type of direct observation.

## One-time surveys

Surveys use a structured questionnaire as the basis for many monitoring and evaluation data collection. Questionnaires allow for focused data collection about specific performance questions or indicators. Questionnaires can provide precise answers to carefully defined questions. Questionnaires can be self administered or administered by interviewers. If questionnaires are self administered, respondents are asked to fill up the questionnaire and return it the M&E staff. Questionnaires can be sent by mail, or electronically. However, non-response rate is higher in self administered questionnaires, although it could save time and cost if effective.

M&E system may need to collect detailed qualitative and quantitative information using structured questionnaires in order to conduct statistical analysis, and generate statistically validated results. In such cases, conducting census on the whole target population of an intervention is usually not feasible and practical. One survey approach is to conduct a one-shot survey on selected samples. In such cases interviews are conducted on appropriately selected sample of the target population. Information gathered from representative samples can then be used to generalize about the target population. It is important to follow appropriate sampling techniques to ensure representativeness of the samples.

## Panel surveys

M&E system may need to collect data and information on a continuous basis in order to monitor changes over time. In such cases, one-shot surveys may not be adequate. Panel surveys are surveys conducted on the same selected sample over time. The same questionnaire is administered at different points in time on the same selected samples. A major problem of panel surveys is that respondents may drop out of the survey for various reasons, including death, moving out of the area, unwillingness to continue in the survey etc. Appropriate panel data analysis can then be used to analyse the data.

## Census

A census is a process of obtaining information about every member of a population. In some case, it may be possible to conduct census on the whole target population of an intervention. However, censuses can be too expensive and time consuming. The quality of census data may also be low, as it takes a lot of human, financial and physical resources and time to make a complete enumeration of each element of the population. Moreover, there may be little advantage of using censuses over sample surveys.

## Field experiments

In rare cases, controlled experiments may be administered in the intervention area. An experiment is a special form of research which sets out to examine the relationship between two factors by manipulating one while measuring changes in the other. For instance, one may plant maize and apply different level of fertilizer to measure the effect of fertilizer on total yield. The advantage of experiments is that it becomes easier to establish cause and effect relationships in analysing impacts. In fact, experimentation is a means of obtaining data with relatively high precision in measurement of the variables. In many instance this precision is associated with a longer time requirement than that needed for obtaining non-experimental data. However, experiments are rarely used in M&E systems.

### 8.3 Choosing data collection method

In trying to choose data collection methods, one should note that there is no correct answer as to which method is the best. It all depends on a given organization's resource availability, M&E information needs, time constraints, and the level of data accuracy and detail required. Hence, in most cases a combination of data collection strategies might work best in building the information system to tracking each indicator. Usually using a combination of different methods is necessary to carry out M&E.

In general, data collection strategies involve some tradeoffs with respect to cost, accuracy, reliability, and timeliness. In this regard, more structured and formal methods of data collection generally tend to be more precise, costly and time consuming. On the other hand, less structured and informal methods are less costly and less time consuming, but they are less precise. If data are needed frequently and on a routine basis to inform management decision-making, it may be preferable to adopt less precise, less structured, and inexpensive data collection strategies.

Triangulation (the use of three or more sources or types of information to verify or substantiate an assessment) is usually an important approach used to check the reliability of data. For instance, a participatory rural appraisal (PRA) process used to find out the benefit of an activity to the users may combine a number of different methods such as transect walks, matrix ranking, focus group discussion etc. Similarly a household survey may combine interviewing, discussion and facilitation methods. For example, the output monitoring activity of the IPMS project used a methodology that encompasses different methods including group discussion with project beneficiaries and community representatives, focus group discussion with staff of governmental organizations (such as BoARD and OoARD) and non-governmental organizations, document review and direct observations.

### 8.4 Preparations for data collection

Once the data source and the methods of data collection are known, the next step would be the actual data collection. However, before the start of the actual data collection on indicators, data collection instruments and forms must be prepared. Questionnaires, checklists and other necessary forms for recording data have to be prepared ahead of time for the different methods of data collection. Data collection personnel such as facilitators, enumerators, and field supervisors need to be recruited and trained. All necessary arrangement should be made with partners who would be involved in the data collection.

Interviewers and facilitators are needed to collect data and to conduct group-based discussion/analysis, respectively. Interviewing and facilitating are two complementing skills which may be affected by age,

gender, educational level and other socio-economic factors. These factors may negatively or positively affect the interviewer's or facilitator's capacity to do interviewing or facilitating. If data is to be collected from rural population, interviewers and facilitators must be those who appreciate rural life. Therefore, selecting people who best fit with the task at hand is essential. Moreover, it is good to make sure that the people who are using the methods are comfortable in using the methods.

### Pre-testing data collection instruments

It is important to pre-test the indicators and their information requirements, and the data collection methods and instruments before moving on to a full scale data collection activity. Pre-testing is a means of verifying whether the indicators are valid, their information requirements are feasible, and the methods and instruments are appropriate. It helps to learn what works and what does not and to avoid major mistakes which may save cost and time. Pre-testing may also improve data quality and help to examine the proposed indicators as they relate to data collection strategies. For example, if every indicator would require costly data collection methods, revision of the indicators may be required. M&E system should choose indicators that will yield the required information at lowest cost as possible. Pre-testing may lead to rationalizing and prioritizing the set of indicators. It is useful if enumerators and data collection supervisors could participate in the pre-test.

## 8.5 Data analysis

Whether we are looking at monitoring or evaluation, at some point, we are going to find ourselves with a large amount of information and we will have to decide how to analyse it. If we are using an external evaluation team, it will be up to this team to do the analysis, but, sometimes in evaluation, and certainly in monitoring, the organization or project/program have to do the analysis.

Thus, once M&E data is collected, it must be analysed, make conclusions and develop recommendations. The analysis method should depend on the type of M&E finding that is sought for and the type of information that is analysed. M&E data can be analysed in various ways. Tabular and graphical analyses are widely used in M&E systems. The most common method of analysis is computing deviations from the baseline and targets, as direct measures of performance. Simple descriptive analysis such as computing means, minimum and maximum values, proportions or percentages, and standard deviations can be computed. Comparison of performance data over time can be very useful. Analysis can be made for subpopulations of the target population in order to understand the impact of the intervention on the different segments of the population. It is important to allow for sufficient time for revision of analysis.

Data analysis should also be guided by the information needs of different users of the M&E findings. In other words, data must be analysed according to the main interests and preferences of each user. In general, there are at least three basic questions that need to be answered before embarking on a specific data analysis exercise. These are:

- i. What is the characteristic of the data at hand: qualitative or quantitative?
- ii. What type of statistical tools is appropriate to achieve the stated objectives and what data structure do the analysis tools require?
- iii. Is there the required expertise to carry out the analysis and interpret the results of the study?

## 8.5.1 Qualitative vs. quantitative data analysis

### Quantitative data analysis

Quantitative analysis demonstrates the degree of achievement or casual relationships using quantitative data based on statistics grounds. It tries to present the M&E results as scientifically as possible. In this regard, different techniques are used. Below we describe briefly some of the techniques.

#### *Simple aggregation*

The basic quantitative analysis method deals with a single variable. This method is suitable for examining the degree of achievement or for comparing that achievement with target values. These types of analysis include computations of frequencies (e.g. the number of persons who answered 'yes' and that of persons who answered 'no'); percentages/proportions (e.g. the ratio of persons who responded per 100 persons); central tendency (the mean, the mode, the median); and standard deviations (to see how far the values are distributed from the mean).

#### *Examination of difference*

This type of analysis helps to identify whether there is significant difference between or among two or more variables. In this regard t-test can be used. This analysis is appropriate when one wants to compare the means of two groups. The one sample t-test procedure tests whether the mean of a single variable differs from a specified constant. On the other hand, the paired samples t-test procedure compares the means of two variables for a single group. It computes the difference between values of the two variables for each case and tests whether the average differs from zero. Analysis of Variance (ANOVA) is another procedure that produces a one-way analysis of variance for quantitative dependent variable by a single independent variable. Analysis of variance is used to test the hypothesis that several means are equal.

#### *Measures of association*

This type of analysis helps to determine statistically whether there exists association between two variables. It is possible to determine the existence of association between two discrete as well as continuous variables. In this regard correlation coefficients and chi-square analysis can be used to see if there are significant associations between variables.

#### *Explaining cause–effect relationship: Regression analysis*

There are different kinds of regression models that can be used to predict the cause–effect relationship between a dependent variable (explained variable) and independent (explanatory) variable(s). The dependent variable is the variable that is explained by the independent variable(s), and is also called the outcome or the effect. The independent variable is the causative agent for the outcome, and is, therefore, called the explanatory variable.

### Qualitative analysis

Qualitative analysis uses qualitative information obtained from literature reviews, interviews, or focus group discussions etc. The set of analytical tools are not prepared in advance. In the process of data analysis, the meaning of data, new facts, or relationships between factors is deductively constructed. The results of qualitative analysis may be influenced by the biases of both respondents and analysis. One of the merits of qualitative analysis is its ability to obtain detailed information about local target populations and people's behavioural changes. The results of qualitative analysis can be used as supporting evidence for the results of quantitative analysis, and thus can help to identify the various factors influencing performance. Below, we describe some of the qualitative analysis methods briefly.

### *Explaining the situation*

Such analysis is used to convey to readers (those who might use the qualitative data) the whole picture of the intervention including what is happening in the area, how stakeholders are perceiving the intervention, and in what situation specific activities or events are being implemented etc.

### *Classify information according to patterns and issues*

Such analysis is used to find out information or the results of observations that can be classified under the same issue or concept and bring them together in a group. Data may not only be labelled, but also classified. It is useful to conduct this task with two or more persons and compare results. This is because different persons may analyse data from different view points, and thus comparing results can reduce the biases of analysts. The classified data can be used to identify the relationships between specific themes and the target intervention.

### *Examine relationships within information*

Another method of qualitative analysis is to examine the relationships within information. The situation and issues of an intervention can be understood by logically classifying qualitative data into such categories as the process and effects of the intervention. Tables or flowcharts may be helpful to identify those categories and explain the relationship among them.



## 9 Reporting and using monitoring and evaluation information

### 9.1 Reporting M&E information

The primary purpose of M&E information is to serve as management tool. Reporting M&E findings is about deciding what is reported, to whom it is reported, and when it is reported. M&E findings can be used for various purposes including:

- demonstrating accountability,
- facilitating organizational learning,
- determining what works and what does not work,
- creating institutional memory through documentation,
- organizing support, and
- creating better understanding of projects/programs/policies.

M&E reporting and communication strategy should consider the intended audience, the format of reporting, the time of reporting, and the delivery mechanism. It may be useful to keep intended audiences up to date during the M&E process to avoid surprises. Continuous communication may be needed for decision-makers. Informal communications such as phone calls, e-mails, fax messages, and formal methods such as briefings, presentations, and written reports should be part of an overall communication strategy.

M&E findings should be used to improve implementation of an intervention and its outcomes. Therefore, the findings should be communicated widely to the relevant stakeholders. The reporting schedule should be determined and stakeholders who should receive the report must be identified. Usually the M&E report is sent to funding agencies for accountability reasons. Implication for policy change must be shared with relevant government officials. Lessons learnt need to be shared with organizational staff and stakeholders.

Data should be presented in an easy-to-understand manner. Only the most important data should be presented. The communication strategy should consider the communication needs of each type of audiences, i.e. the report should be packaged and prepared according to the main needs of specific audiences. It is always useful to report against baselines and targets to determine whether progress has been achieved and sustained.

Comparing actual outcomes to targets is also central in reporting M&E findings, because the report should clearly indicate the status of the intervention with regard to results expectations. Findings and recommendations should be organized around major outcomes and their indicators. M&E findings should also provide significant information regarding trends and directions of the intervention over time. Clues to problems that arise during the course of implementation, and possible ways of making necessary improvements in implementation strategies should also be indicated because decision-makers may be looking for indicators of actions required to improve effectiveness and impact of interventions. It may be important to highlight the implications of actions, since decision-makers usually will want to have fuller understanding of the consequences of their actions. Background information should be limited to the minimum.

The M&E reporting can be done in three major ways: written reports, executive summaries and oral presentations (Kusek and Rist 2004).

### *Written reports*

Written reports should normally contain: (1) Introduction (background to intervention, context of intervention area, purpose of report, M&E questions, and information about goals and objectives), (2) Methodology of the M&E (M&E focus, data and data sources, data analysis, when and by whom the M&E was conducted, and limitations of the methodology), (3) Key findings organized around major outcomes and their indicators, and (4) Conclusions and recommendations closely connected with findings. Written reports can include tables, charts, graphs and other data and result presentation formats.

### *Executive summaries*

Executive summaries are usually short (one to four pages). Executive summaries contain brief background and purpose of the M&E, brief description of the evaluations questions and the methods, and major findings and recommendations. In executive summaries major findings and recommendations can be presented in bullet point format. Readers of the executive summary can be referred to the major report for further information.

### *Oral presentations*

Oral presentations can be used either alone or in combination with written reports. Oral presentations should be simple, clear, and tailored to the intended audience. Whenever possible, oral presentations can be made interactive with the audience. Oral presentations can make use of visual aids such as tables, charts, graphs and maps. Visuals can illustrate directions and trends at a glance. Tables are best used to present data, highlighting changes, comparisons and relationships.

Apart from sending the report, arranging sessions/meetings for critical reflection by stakeholders on the findings of the M&E report is important in order to get a clearer and common understanding of the M&E findings, and facilitate the use of the findings. For example, in the IPMS project, the formal annual joint RALC/WALC meetings and field visits provide relevant stakeholders the opportunity to reflect on the progress of the project/program.

## 9.2 Using M&E information

M&E information has a wide application. Kusek and Rist (2004) identified the following uses of performance feedbacks:

- Make corrective decisions to improve performance of organizations or interventions
- Enhance organizational learning and build institutional memory
- Report for accountability
- Formulate and justify budgets
- Make resource allocation decisions
- Motivate personnel to make program improvements
- Formulate and monitor the performance of contractors and grantees
- Provide data for special, in-depth evaluation
- Support strategic and other long-term planning efforts and
- Communicate better with the public to build trust.

Continuous assessment and evaluation of performance and feedback enhances learning and builds institutional memory. The loss of institutional memory that could occur due to staff turnover can be minimized and changing management becomes easier with an M&E system in place. Knowledge and knowledge management are important components of using performance findings. Knowledge

management in M&E implies capturing findings and insights, institutionalizing learning, and organizing the wealth of information produced continually by the M&E system. Project/program/policy evaluations should be means of systematic organizational learning, rather than one-shot or *ad hoc* events.

Accountability and transparency are being increasingly demanded by public agencies. Timely reporting of performance of interventions to the public builds public trust and generates support. Budgetary allocations of projects/programs/policies can be increased or reduced depending on the M&E findings and recommendations. In some cases, interventions may be eliminated fully. Human beings respond to incentives. M&E findings can be used to reward and motivate personnel for good performance or sanction for sub performance. Progress of contracts need to be monitored before additional funds are released or contracts are renewed. M&E, as an information system, builds databases, which can be used for further analysis with or without additional data. M&E findings and databases can also be useful inputs for planning.

In order to support strategic and other long-term planning efforts, the M&E data that has been collected has to be stored and managed in order to facilitate accessibility over time. To ensure accessibility for future users, it must be stored and described in suitable ways, i.e. metadata needs to be developed. Metadata is information about the data—such as what, how, where, when, how often and by whom data was collected, as well as how it is recorded, stored and analysed. Metadata helps to determine whether or not data sets collected at different time periods are compatible and so able to be combined to build time series data. Therefore, metadata acts like a library catalogue. In this way it assists the public, other agencies, and own staff to locate all available data in a field of interest. This helps to prevent duplication of effort and to share knowledge.

# 10 Institutionalizing and sustaining the RBM&E system

## 10.1 Institutionalizing RBM&E

Institutionalizing RBM&E in the systems and structures of organizations dealing with agricultural development is critical to their ability to promote agricultural development. Institutionalization of RBM&E means creating a RBM&E system with policy, legal and institutional arrangements to produce monitoring information and evaluation findings which are judged valuable by key stakeholders (Sivagnanasothy 2007). When RBM&E is institutionalized, it serves as an integral part of the development project/program/policy cycle to improve performance and accountability, to provide effective feedback, to improve planning, budgeting and policymaking to achieve development effectiveness. In order to institutionalize RBM&E, building skills of staff, developing procedures, methodology, data systems, manuals etc. are important issues that need to be considered. Dissemination mechanisms of RBM&E findings also need to be put in place so that formalized feedback arrangements operate to integrate lessons into the planning and design of new projects/programs/policies. Institutionalizing RBM&E system also calls for the establishment of strong links between RBM&E, and policy formulation, reforms, planning, budgeting and resource allocation functions (Mackay 2007; Sivagnanasothy 2007).

Sivagnanasothy (2007) pointed out the following issues as important strategies that aid effective institutionalization of RBM&E system:

- Providing policy commitment and support
- Providing legal and budgetary support
- Providing sound institutional arrangement (i.e. ensure proper institutional arrangements to place evaluation in a strategic context)
- Strengthening methodologies and practices
- Building evaluation capacity of staff
- Creating/strengthening feedback and information dissemination mechanisms.

## 10.2 Sustaining RBM&E

Sustainability and use of RBM&E systems are interdependent. Systems that are not used will not be sustainable. The issue of use has to be addressed first because it is the prerequisite to system sustainability. Building a RBM&E system should be regarded as a long-term effort, as opposed to a periodic effort for a short period or for the duration of a specific project/program/policy. Sustaining such systems within governments or organizations recognizes the long term process involved in ensuring utility.

Kusek and Rist (2004) identified six critical components of sustaining a RBM&E: demand, clear roles and responsibilities, trustworthy and credible information, accountability, capacity and incentives. Each of these components needs continuous attention. Below, we give brief description of each of these critical components.

### *Demand*

If demand for RBM&E information is periodic, RBM&E systems are not going to be used and sustained. Structured requirements for reporting results can help lead to sustained and consistent demand for such systems. In many cases, demand can also be stimulated when the strategic goals are translated into RBM&E systems.

### *Clear roles and responsibilities*

Clear roles, responsibilities and formal organizational lines of authority must be established. The unit and people who will be in charge of collecting, analysing, and reporting performance information must be clearly defined.

### *Trustworthy and credible information*

The RBM&E system must be able to produce information that brings both good and bad news. Performance information should be transparent and made available to all key stakeholders. If debate on issues is not backed up by trustworthy and credible information, only personal opinions and presumptions are left. It should also be noted that the producers of information need protection from political reprisals. If bad news brings career problems to the messengers, fear will permeate the system and the reliability of the information produced will be compromised.

### *Accountability*

Accountability means that problems should be acknowledged and addressed. No part of the government should be exempt from accountability to stakeholders. Civil society organizations and NGOs can play role in encouraging transparency and accountability.

### *Capacity*

Sound technical skills in data collection and analysis are necessary for the system's sustainability. Managerial skills in strategic goal setting and organizational development are also needed. Data collection and retrieval systems must be up and running, and modernized. Organizations will need to commit continuous financial resources to the upkeep and management of RBM&E systems. Institutional experience and memory are also helpful in the long-term sustainability of these systems.

### *Incentives*

Incentives need to be introduced to encourage use of performance information. This means that success needs to be acknowledged and rewarded, problems need to be addressed, organizational learning is valued, and budget savings are shared. Corrupt or ineffective systems cannot be counted on to produce quality information and analysis.

## 10.3 The importance of incentives and disincentives in sustaining RBM&E systems

Sustaining RBM&E systems also involves using appropriate incentives to keep managers and stakeholders on track and motivated. There are a variety of organizational, financial, resources, political, technical assistance, and training incentives that can be used to sustain RBM&E systems. Likewise, managers need to remove disincentives to sustaining RBM&E systems. Boxes 3 and 4 contain checklists of the kinds of incentives and disincentives that should be considered, respectively.

### Box 3: Checklist for staff incentives that encourage learning-oriented, RBM&E

Are the following incentives in place?

- Clarity of RBM&E responsibility
- Financial and other rewards: appropriate salaries and other rewards
- Activity support: support, such as financial and other resources, for carrying out RBM&E activities
- Personnel and partner strategy: hiring staffs that have an open attitude to learning, and signing on partners who are willing to try out more participatory forms of RBM&E
- Project/program/policy culture: compliments and encouragements for those who ask questions and innovate, giving relatively high status to RBM&E among staff
- Performance appraisal processes: equal focus on staff capacity to learn and innovate, rather than focusing only on achievement of quantitative targets
- Showing the use of RBM&E data: making the data explicit and interesting by displaying them
- Feedback: telling data collectors, information providers, and others involved in the process how their data was used (analysed), and what it contributed to the project.

Source: IFAD cited by Kusek and Rist (2004).

### Box 4: Checklist for staff disincentives that hinder learning-oriented, RBM&E

Have the following disincentives been removed from project/program/policy?

- Using the RBM&E unit as the place to park demoted or unqualified staff
- Not making clear how data will be or were used
- Chastising those who innovate within their project boundaries or those who make mistakes
- Focusing performance appraisals only on activities undertaken
- Frequent rotation of staff to different posts
- Staff feeling isolated or helpless in terms of their contribution being recognized toward achieving the project/program goal (the 'line of sight' issue)
- Unconstructive attitudes toward what constitutes participation or toward the primary stakeholder groups.

Source: IFAD as cited by Kusek and Rist (2004).

## 10.4 Challenges in institutionalizing and sustaining RBM&E systems

There are a number of challenges that may arise in institutionalizing and sustaining RBM&E systems (Kusek and Rist 2004). Some of the most critical challenges in implementing and sustaining RBM&E systems are the challenges in the human resource area. These challenges are perhaps not so different from all public sector human resource matters, but there are unique dimensions that have to be

addressed. First, there are issues in recruiting and holding talented staff who can build and manage a new information system. Can they be found and, if so, can they be hired? Second is the issue of what staff will risk venturing into a new government initiative or stated differently, what is the calibre of those who leave their present positions for positions in a new RBM&E unit? Third is the matter of whether the first cohorts of those hired are change agents. Building RBM&E system is a politically charged change process. Do those being hired understand this and are they ready to manage a change process? Fourth is whether continuous training can be provided for all personnel at all levels. New methodologies, technologies, and procedures are inevitable and need to be shared with staff. Furthermore, given staff turnover, how soon and how adequately can new staff be trained to quickly increase their productivity and contributions to the unit?

The RBM&E system will have to respond and adapt to changes in legislative and organizational priorities. In spite of these larger political and environmental changes, maintaining indicator stability overtime is important because one may want to be able to compare similar issues and trends over a given period of time. Generally, RBM&E systems are essentially political challenges, and to a lesser extent, technical ones. Creating, implementing, and sustaining RBM&E systems can help to bring about major cultural changes in the way governments and organizations operate. RBM&E systems can bring about positive cultural changes that lead to improved performance, enhanced accountability and transparency, and learning and knowledge.

## References

- Alur A, Snehlata N and Prem K. 2005. *Participatory monitoring and evaluation-field experiences*. NGO Programme Karnataka-Tamil Nadu Series 1. Intercooperation; Delegation, Hyderabad, India.
- Anandajayasekeram P, van Rooyen CJ, Rukuni M, Marassas C and D'Haese M. 2004. *Agricultural project planning and analysis: A source book*. Second edition. A joint publication by University of Pretoria, FARMESA and University of Ghent, Gent, Belgium.
- Catley A, Burns J, Abebe D and Suji O. 2008. *Participatory impact assessment: A guide for practitioners*. Feinstein International Center, Tufts University, Medford, USA.
- CIDA (Canadian International Development Agency). 1999. *Results-based management in CIDA: An introductory guide to the concepts and principles*. CIDA, Ottawa, Canada.
- CIDA (Canadian International Development Agency). 2008a. Presentation of the amended key results-based management terms and definitions for the 2008 results-based management policy statement. Performance Management Division, Strategic Policy and Performance Branch. CIDA, Ottawa, Canada.
- CIDA (Canadian International Development Agency). 2008b. *Results-based management policy statement*. CIDA, Ottawa, Canada. Available at <http://www.acdi-cida.gc.ca/CIDAWEB/acdicida.nsf/En/ANN-102084042-GVJ#PDF>. Last updated August 8<sup>th</sup>, 2008.
- CIDA (Canadian International Development Agency). 2009. *RBM tools at CIDA: How to guide. Instruction for the use of CIDA's three main results-based management (RBM) working tools; the logic model, performance measurement framework and risk register*. Last updated 06/04/09. CIDA, Ottawa, Canada.
- DAC Working Party on Aid Evaluation. 2000. *Result based management in the development co-operation agencies: A review of experience (Executive Summary)*. Available at [http://www.oecd.org/secure/pdfDocument/0,2834,en\\_21571361\\_34047972\\_31950682\\_1\\_1\\_1\\_1,00.pdf](http://www.oecd.org/secure/pdfDocument/0,2834,en_21571361_34047972_31950682_1_1_1_1,00.pdf).
- Earl S, Carden F and Smutylo T. 2001. *Outcome mapping: Building learning and reflection into development programs*. IDRC (International Development Research Centre), Ottawa, Canada.
- Elliott J, Heesterbeek S, Lukensmeyer CJ and Slocum N. 2005. *Participatory methods toolkit: A practitioner's manual*. The King Baudouin Foundation and the Flemish Institute for Science and Technology Assessment (viWTA). Available at <http://www.viWTA.be/files/Toolkit%20deliberative%20polling.pdf>.
- Estrella M and Gaventa J. 1998. *Who counts reality? Participatory monitoring and evaluation: A literature review*. IDS Working Paper No. 70. University of Sussex, Brighton, UK.
- Friis-Hansen E and Sthapit B. (eds). 2000. *Participatory approaches to the conservation and use of plant genetic resources*. IPGRI (International Plant Genetic Resources Institute), Rome, Italy. 216 pp.
- IFAD (International Fund for Agricultural Development). 2002. *A guide for project monitoring and evaluation: Managing for impact in rural development*. IFAD, Rome, Italy. Available at <http://www.ifad.org/evaluation/guide/>.
- Kusek JZ and Rist R. 2004. *Ten steps to a results-based monitoring and evaluation system. A handbook for development practitioners*. The World Bank, Washington, DC, USA.
- Mackay K. 2007. *How to build monitoring and evaluation systems to support better government*. World Bank Independent Evaluation Group, Washington, DC, USA.
- Narayan-Parker D. 1993. *Participatory evaluation: Tools for managing change in water and sanitation*. World Bank Technical Paper 207. World Bank, Washington, DC, USA.
- OECD (Organization for Economic Cooperation and Development). 2002. *Glossary of key terms in evaluation and results-based management*. OECD/DAC, Paris, France.
- Roche C. 1999. *Impact assessment for development agencies: Learning to value change*. Oxfam publications, Oxford, UK.
- Rossi PH, Lipsey MW and Freeman HE. 2004. *Evaluation: A systematic approach*. 7th ed. Sage Publications, Inc, California, USA.
- Sivagnanasothy V. 2007. *Institutionalization of monitoring and evaluation system in Sri Lanka—Lessons learnt, best practices, issues, challenges and the way forward*. Department of Foreign Aid and Budget Monitoring, Ministry of Plan Implementation, Sri Lanka.
- UNDP (United Nations Development Program). 2002. *Handbook on monitoring and evaluating for results*. UNDP Evaluation Office, New York, USA.
- UNICEF and MSH. 1998. *The guide to managing for quality*. Available at <http://erc.msh.org/quality/directob.cfm>.
- USAID. 1996. *Establishing performance targets. Performance monitoring and evaluation*. TIPS Number 8. Available at [http://www.dec.org/pdf\\_docs/pnaaby226.pdf](http://www.dec.org/pdf_docs/pnaaby226.pdf).



## Annex 1 Glossary of common M&E terms (OECD 2002)

This glossary is taken from the 2002 OECD/DAC glossary of key terms in evaluation and results based management. Although some of the terminologies are not used in the guide, we have included the glossary for reference purpose.

### A

*Accountability:* Obligation to demonstrate that work has been conducted in compliance with agreed rules and standards or to report fairly and accurately on performance results *vis-à-vis* mandated roles and/or plans. This may require a careful, even legally defensible, demonstration that the work is consistent with the contract terms.

Note: Accountability in development may refer to the obligations of partners to act according to clearly defined responsibilities, roles and performance expectations, often with respect to the prudent use of resources. For evaluators, it connotes the responsibility to provide accurate, fair and credible monitoring reports and performance assessments. For public sector managers and policymakers, accountability is to taxpayers/citizens.

*Activity:* Actions taken or work performed through which inputs, such as funds, technical assistance and other types of resources are mobilized to produce specific outputs.

*Analytical tools:* Methods used to process and interpret information during an evaluation.

*Appraisal:* An overall assessment of the relevance, feasibility and potential sustainability of a development intervention prior to a decision of funding.

*Assumptions:* Hypotheses about factors or risks which could affect the progress or success of a development intervention.

Note: Assumptions can also be understood as hypothesized conditions that bear on the validity of the evaluation itself, e.g. about the characteristics of the population when designing a sampling procedure for a survey. Assumptions are made explicit in theory based evaluations where evaluation tracks systematically the anticipated results chain.

*Attribution:* The ascription of a causal link between observed (or expected to be observed) changes and a specific intervention.

Note: Attribution refers to that which is to be credited for the observed changes or results achieved. It represents the extent to which observed development effects can be attributed to a specific intervention or to the performance of one or more partner taking account of other interventions, (anticipated or unanticipated) confounding factors, or external shocks.

*Audit:* An independent, objective assurance activity designed to add value and improve an organization's operations. It helps an organization accomplish its objectives by bringing a systematic, disciplined approach to assess and improve the effectiveness of risk management, control and governance processes.

Note: a distinction is made between regularity (financial) auditing, which focuses on compliance with applicable statutes and regulations; and performance auditing, which is concerned with relevance, economy, efficiency and effectiveness. Internal auditing provides an assessment of internal controls

undertaken by a unit reporting to management while external auditing is conducted by an independent organization.

## B

*Base-line study:* An analysis describing the situation prior to a development intervention, against which progress can be assessed or comparisons made.

*Benchmark:* Reference point or standard against which performance or achievements can be assessed.

Note: A benchmark refers to the performance that has been achieved in the recent past by other comparable organizations, or what can be reasonably inferred to have been achieved in the circumstances.

*Beneficiaries:* The individuals, groups, or organizations, whether targeted or not, that benefit, directly or indirectly, from the development intervention.

## C

*Cluster evaluation:* An evaluation of a set of related activities, projects and/or programs.

*Conclusion:* Conclusions point out the factors of success and failure of the evaluated intervention, with special attention paid to the intended and unintended results and impacts, and more generally to any other strength or weakness. A conclusion draws on data collection and analyses undertaken, through a transparent chain of arguments.

*Counterfactual:* The situation or condition which hypothetically may prevail for individuals, organizations, or groups where there is no development intervention.

*Country program evaluation/Country assistance evaluation:* Evaluation of one or more donor's or agency's portfolio of development interventions, and the assistance strategy behind them, in a partner country.

## D

*Data collection tools:* Methodologies used to identify information sources and collect information during an evaluation.

Note: Examples are informal and formal surveys, direct and participatory observation, community interviews, focus groups, expert opinion, case studies, literature search.

*Development intervention:* An instrument for partner (donor and non-donor) support aimed to promote development.

Note: Examples are policy advice, projects, programs.

*Development objective:* Intended impact contributing to physical, financial, institutional, social, environmental, or other benefits to a society, community, or group of people via one or more development interventions.

## E

*Economy:* Absence of waste for a given output.

Note: An activity is economical when the costs of the scarce resources used approximate the minimum needed to achieve planned objectives.

*Effect:* Intended or unintended change due directly or indirectly to an intervention. Related terms: results, outcome.

*Effectiveness:* The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.

Note: Also used as an aggregate measure of (or judgement about) the merit or worth of an activity, i.e. the extent to which an intervention has attained, or is expected to attain, its major relevant objectives efficiently in a sustainable fashion and with a positive institutional development impact. Related term: efficacy.

*Efficiency:* A measure of how economically resources/inputs (funds, expertise, time etc.) are converted to results.

*Evaluability:* Extent to which an activity or a program can be evaluated in a reliable and credible fashion.

Note: Evaluability assessment calls for the early review of a proposed activity in order to ascertain whether its objectives are adequately defined and its results verifiable.

*Evaluation:* The systematic and objective assessment of an ongoing or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors. Evaluation also refers to the process of determining the worth or significance of an activity, policy or program. An assessment, as systematic and objective as possible, of a planned, ongoing, or completed development intervention.

Note: Evaluation in some instances involves the definition of appropriate standards, the examination of performance against those standards, an assessment of actual and expected results and the identification of relevant lessons. Related term: review.

*Ex-ante evaluation:* An evaluation that is performed before implementation of a development intervention. Related terms: appraisal, quality at entry.

*Ex-post evaluation:* Evaluation of a development intervention after it has been completed.

Note: It may be undertaken directly after or long after completion. The intention is to identify the factors of success or failure, to assess the sustainability of results and impacts, and to draw conclusions that may inform other interventions.

*External evaluation:* The evaluation of a development intervention conducted by entities and/or individuals outside the donor and implementing organizations.

## F

*Feedback:* The transmission of findings generated through the evaluation process to parties for whom it is relevant and useful so as to facilitate learning. This may involve the collection and dissemination of findings, conclusions, recommendations and lessons from experience.

*Finding:* A finding uses evidence from one or more evaluations to allow for a factual statement.

*Formative evaluation:* Evaluation intended to improve performance, most often conducted during the implementation phase of projects or programs.

Note: Formative evaluations may also be conducted for other reasons such as compliance, legal requirements or as part of a larger evaluation initiative.

## G

*Goal:* The higher-order objective to which a development intervention is intended to contribute. Related term: development objective.

## I

*Impacts:* Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.

*Independent evaluation:* An evaluation carried out by entities and persons free of the control of those responsible for the design and implementation of the development intervention.

Note: The credibility of an evaluation depends in part on how independently it has been carried out. Independence implies freedom from political influence and organizational pressure. It is characterized by full access to information and by full autonomy in carrying out investigations and reporting findings.

*Indicator:* Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor.

*Inputs:* The financial, human, and material resources used for the development intervention.

*Institutional development impact:* The extent to which an intervention improves or weakens the ability of a country or region to make more efficient, equitable, and sustainable use of its human, financial, and natural resources, for example through:

(a) better definition, stability, transparency, enforceability and predictability of institutional arrangements and/or

(b) better alignment of the mission and capacity of an organization with its mandate, which derives from these institutional arrangements. Such impacts can include intended and unintended effects of an action.

*Internal evaluation:* Evaluation of a development intervention conducted by a unit and/or individuals reporting to the management of the donor, partner, or implementing organization.

## J

*Joint evaluation:* An evaluation to which different donor agencies and/or partners participate.

Note: There are various degrees of 'jointness' depending on the extent to which individual partners cooperate in the evaluation process, merge their evaluation resources and combine their evaluation reporting. Joint evaluations can help overcome attribution problems in assessing the effectiveness of programs and strategies, the complementarity of efforts supported by different partners, the quality of aid coordination etc.

## L

*Lessons learned:* Generalizations based on evaluation experiences with projects, programs, or policies that abstract from the specific circumstances to broader situations. Frequently, lessons highlight strengths or weaknesses in preparation, design, and implementation that affect performance, outcome, and impact.

*Logical framework (Logframe):* Management tool used to improve the design of interventions, most often at the project level. It involves identifying strategic elements (inputs, outputs and outcomes) and their causal relationships, indicators, and the assumptions or risks that may influence success and failure. It thus facilitates planning, execution and evaluation of a development intervention.

## M

*Meta-evaluation:* The term is used for evaluations designed to aggregate findings from a series of evaluations. It can also be used to denote the evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators.

*Mid-term evaluation:* Evaluation performed towards the middle of the period of implementation of the intervention.

*Monitoring:* A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.

## O

*Outcome:* The likely or achieved short-term, medium-term effects of an intervention's outputs.

*Note:* According to CIDA 2008b new RBM policy statement outcomes are the same as results, and are further qualified as immediate (short-term), intermediate (medium-term), or ultimate (long-term). Immediate outcome is a change that is directly attributable to the outputs of an organization, policy, program, or initiative. Intermediate outcome is a change that is expected to logically occur once one or more immediate outcomes have been achieved. Ultimate outcome is the highest level change that can be reasonably attributed to an organization, policy, program, or initiative in a casual manner, and is the consequence of one or more intermediate outcomes (CIDA 2009).

*Outputs:* The products, capital goods and services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes.

*Note:* According to CIDA 2008 RBM policy statement, output is defined as direct products or services stemming from the activities of an organization, policy, program or initiative (CIDA 2009).

## P

*Participatory evaluation:* Evaluation method in which representatives of agencies and stakeholders (including beneficiaries) work together in designing, carrying out and interpreting an evaluation.

*Partners:* The individuals and/or organizations that collaborate to achieve mutually agreed upon objectives.

Note: The concept of partnership connotes shared goals, common responsibility for outcomes, distinct accountabilities and reciprocal obligations. Partners may include governments, civil society, non-governmental organizations, universities, professional and business associations, multilateral organizations, private companies etc.

*Performance:* The degree to which a development intervention or a development partner operates according to specific criteria/standards/guidelines or achieves results in accordance with stated goals or plans.

*Performance indicator:* A variable that allows the verification of changes in the development intervention or shows results relative to what was planned.

*Performance measurement:* A system for assessing performance of development interventions against stated goals.

*Performance monitoring:* A continuous process of collecting and analysing data to compare how well a project, program, or policy is being implemented against expected results.

*Process evaluation:* An evaluation of the internal dynamics of implementing organizations, their policy instruments, their service delivery mechanisms, their management practices, and the linkages among these.

*Program evaluation:* Evaluation of a set of interventions, marshalled to attain specific global, regional, country, or sector development objectives.

Note: A development program is a time bound intervention involving multiple activities that may cut across sectors, themes and/or geographic areas.

*Project evaluation:* Evaluation of an individual development intervention designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader program.

Note: Cost–benefit analysis is a major instrument of project evaluation for projects with measurable benefits. When benefits cannot be quantified, cost effectiveness is a suitable approach.

*Project or program objective:* The intended physical, financial, institutional, social, environmental, or other development results to which a project or program is expected to contribute.

*Purpose:* The publicly stated objectives of the development program or project.

## Q

*Quality assurance:* Quality assurance encompasses any activity that is concerned with assessing and improving the merit or the worth of a development intervention or its compliance with given standards.

Note: examples of quality assurance activities include appraisal, RBM, reviews during implementation, evaluations etc.

Quality assurance may also refer to the assessment of the quality of a portfolio and its development effectiveness.

## R

*Reach:* The beneficiaries and other stakeholders of a development intervention.

*Recommendations:* Proposals aimed at enhancing the effectiveness, quality, or efficiency of a development intervention; at redesigning the objectives; and/or at the reallocation of resources. Recommendations should be linked to conclusions.

*Relevance:* The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities and partners' and donors' policies.

Note: Retrospectively, the question of relevance often becomes a question as to whether the objectives of an intervention or its design are still appropriate given changed circumstances.

*Reliability:* Consistency or dependability of data and evaluation judgements, with reference to the quality of the instruments, procedures and analyses used to collect and interpret evaluation data.

Note: Evaluation information is reliable when repeated observations using similar instruments under similar conditions produce similar results.

*Results:* The output, outcome or impact (intended or unintended, positive and/or negative) of a development intervention.

Note: According to CIDA 2008b RBM policy a result is defined as a describable or measurable change that is derived from a cause and effect relationship, and results are the same as outcomes which are further qualified as immediate (short term), intermediate (medium term), or ultimate (long term) (CIDA 2009).

*Results chain:* The causal sequence for a development intervention that stipulates the necessary sequence to achieve desired objectives—beginning with inputs, moving through activities and outputs, and culminating in outcomes, impacts, and feedback. In some agencies, reach is part of the results chain.

Note: At CIDA a results chain (also called a Logic Model) is a depiction of the casual or logical relationships between inputs, activities, outputs and the outcomes of a given policy, program or investment (CIDA 2009).

*Results framework:* The program logic that explains how the development objective is to be achieved, including causal relationships and underlying assumptions.

*Results-Based Management (RBM):* A management strategy focusing on performance and achievement of outputs, outcomes and impacts.

Note: At CIDA RBM is described as 'a life-cycle approach to management that integrates strategy, people, resources, processes and measurements to improve decision-making, transparency, and accountability.'

*Review:* An assessment of the performance of an intervention, periodically or on an *ad hoc* basis.

Note: Frequently 'evaluation' is used for a more comprehensive and/or more in-depth assessment than 'review'. Reviews tend to emphasize operational aspects. Sometimes the terms 'review' and 'evaluation' are used as synonyms.

*Risk analysis:* An analysis or an assessment of factors (called assumptions in the logframe) affect or are likely to affect the successful achievement of an intervention's objectives. A detailed examination of the potential unwanted and negative consequences to human life, health, property, or the environment posed by development interventions; a systematic process to provide information regarding such undesirable consequences; the process of quantification of the probabilities and expected impacts for identified risks.

## S

*Sector program evaluation:* Evaluation of a cluster of development interventions in a sector within one country or across countries, all of which contribute to the achievement of a specific development goal.

Note: A sector includes development activities commonly grouped together for the purpose of public action such as health, education, agriculture, transport etc.

*Self-evaluation:* An evaluation by those who are entrusted with the design and delivery of a development intervention.

*Stakeholders:* Agencies, organizations, groups or individuals who have a direct or indirect interest in the development intervention or its evaluation.

*Summative evaluation:* A study conducted at the end of an intervention (or a phase of that intervention) to determine the extent to which anticipated outcomes were produced. Summative evaluation is intended to provide information about the worth of the program.

*Sustainability:* The continuation of benefits from a development intervention after major development assistance has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time.

## T

*Target group:* The specific individuals or organizations for whose benefit the development intervention is undertaken.

*Terms of reference:* Written document presenting the purpose and scope of the evaluation, the methods to be used, the standard against which performance is to be assessed or analyses are to be conducted, the resources and time allocated, and reporting requirements. Two other expressions sometimes used with the same meaning are 'scope of work' and 'evaluation mandate'.

*Thematic evaluation:* Evaluation of a selection of development interventions, all of which address a specific development priority that cuts across countries, regions, and sectors.

*Triangulation:* The use of three or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment.

Note: By combining multiple data-sources, methods, analyses or theories, evaluators seek to overcome the bias that comes from single informants, single-methods, single observer or single theory studies.

## V

*Validity:* The extent to which the data collection strategies and instruments measure what they purport to measure.





ISBN 92-9146-245-4



በኢትዮጵያ ፌዴራላዊ ዲሞክራሲያዊ ሪፐብሊክ  
የግብርናና ገጠር ልማት ሚኒስቴር

Federal Democratic Republic of Ethiopia  
MINISTRY OF AGRICULTURE AND  
RURAL DEVELOPMENT



Canadian International  
Development Agency

Agence canadienne de  
développement international