

PREFACE

Having the capacity to track results and to use that knowledge to learn what does and what does not work – or how to make things work better – makes M&E a powerful tool for improving development processes and outcomes. In 2006, the Global Donor Platform for Rural Development (GDPRD) and the World Bank undertook to prepare this Sourcebook in collaboration with the Food and Agriculture Organization of the United Nations (FAO). The Sourcebook develops a framework for standardizing approaches for selecting indicators and proposes a menu of core indicators for monitoring and evaluating agriculture and rural development (ARD) activities. Ultimately, the objective is to improve the quality of monitoring and evaluation of agriculture and rural development programmes at the national and global levels.

M&E is intrinsically challenging and requires a level of technical capacity often unavailable in developing countries. The challenge is greater in the poorest countries and in post-conflict situations where less-than-optimal conditions, in particular, the weak statistical capacity, can cause major difficulties. This Sourcebook provides guidance on how to build the capacity needed for effective M&E in developing countries, starting with the identification and collection of the indicators. It suggests a number of approaches for determining which indicators to select given the different types of information that are most pertinent to different agricultural and rural activities, projects and programmes, and data availability. In addition, an innovative feature of the Sourcebook is the presentation of a core set of standard ARD indicators, with the recommendation that they should be regularly compiled by all countries. These “priority indicators” should be the same in all countries so as to allow for country comparisons, and to facilitate the monitoring of ARD programmes and goals at the international level. The Sourcebook identifies a core list of 19 priority indicators, as well as a menu of some 86 indicators that are categorized by sector, subsector and theme. It is hoped that countries may refer to and borrow from it when developing their own national ARD M&E programme. The menu of indicators was validated through in-country workshops in Cambodia, Nicaragua, Nigeria, Senegal and the United Republic of Tanzania.

This Sourcebook was prepared by a team of staff from the World Bank and FAO. Other member institutions of the GDPRD provided valuable inputs. Their remarks, as well as the analysis presented herein, will inform the ongoing GDPRD-facilitated dialogue among donors and partner governments on how to utilize statistics data to improve the management of agriculture, and to capitalize

on its special qualities as a high impact sector with regard to poverty reduction. The recommendations presented in this Sourcebook will also be applied in the *Code of Conduct for More Effective Agriculture and Rural Development Programmes* currently being developed by the GDPRD members.

The aid effectiveness agenda has put considerable pressure on all sectors to empirically demonstrate their performance. It is hoped that this Sourcebook will build upon practitioners' capacity to validate the effectiveness and impacts of agricultural and rural operations.

Christoph Kohlmeyer

Chair
Global Donor Platform
for Rural
Development

Juergen Voegele

Director
Agriculture and Rural
Development Department
World Bank

Hafez Ghanem

Assistant Director General
Economic and Social
Development Department
FAO

EXECUTIVE SUMMARY

BACKGROUND

At the United Nations Conference on Financing for Development, held in Monterrey, Mexico in 2002, both developing and developed countries made commitments to a shared responsibility to achieve development results, particularly those embodied in the Millennium Development Goals. Emphasizing results-based development requires the capacity to monitor indicators that reliably reflect results at all stages of the development process, from strategic planning to implementation to completion. Yet, donors and development practitioners still lack a common framework of results indicators to measure the effectiveness of development assistance. Developing a Monitoring and Evaluation (M&E) system that tracks these indicators using accurate and timely data is therefore a natural priority for the international development community as well as for developing countries themselves. For agencies and institutions involved in agriculture and rural development (ARD), this means developing a common framework that will enable donor agencies to harmonize their monitoring activities.

The reality is that many countries lack the capacity to produce and report the data necessary to inform the international development debate or to monitor their national trends. Although the situation is improving, global databases are still suffering from data gaps and inconsistencies as a result of weaknesses in National Statistical Systems (NSSs). In the final analysis, the validity of global monitoring systems depends on the quality of the data that comes from the countries. It is at the country level that problems occur, and it is at this level that assistance is required to build up sustainable capacity to collect and disseminate appropriate indicators.

DEFINITION, OBJECTIVES AND METHODOLOGY

Monitoring and evaluation are separate but closely connected activities. Monitoring is generally defined as a continuing activity that involves the collection of data on a regular, ongoing basis in order to track inputs, outputs, outcomes and impact while the project/programme is being executed. Evaluation, on the other hand, may use monitoring data, but is carried out at distinct and discreet moments of time to determine the worth or significance of a development activity, policy or programme. Taken together, they form a powerful instrument for planning the future on the basis of what can be shown to work and what does not.

Strengthening capacity for M&E at the subnational and national levels is intrinsically linked to M&E at the global level. Both depend on sound indicators

based on reliable and more complete data. To this end, the Global Donor Platform for Rural Development (GRPRD), the World Bank and the FAO set out to develop a menu of core indicators that could be used to monitor ARD at the project, national, regional and global levels. The approach is generic, but specific indicators are suggested that allow comparisons to be made **between** urban and rural areas, as well as **within** rural areas, specifically between agriculture- and non-agriculture-dependant communities and households. Separate sets of indicators are suggested for: the ARD sector as a whole; various subsectors (crops, livestock, forestry, fisheries and aquaculture, rural micro and small and medium-sized enterprise (SME) finance, research and extension, irrigation and drainage, agribusiness and market development); and related thematic areas (community-based rural development, natural resource management, and agricultural policies and institutions).

The purpose of this Sourcebook is to pull together into a single document a collection of common sense tips and recommendations based on actual practices and experience around the world. The Sourcebook aims first and foremost to help strengthen M&E capacity at the national and subnational levels, and to ensure a consistency of approach and methodology so that, at the global level, sufficient reliable and timely information can be accessed from the different countries and used to make cross-country comparisons and to calculate development indicators at the global level.

The ideal environment for establishing a good M&E system is where: (i) there is a strong and consistent demand for information; (ii) the concept of “management by results” is widely practised; (iii) timely and relevant information is systematically used to improve decision-making and to advance the process of development; and (iv) systems are in place to ensure that reliable and relevant information is available when needed. The less-than-ideal situation, on the other hand, is where (i) demand is weak; (ii) evidence is not used to inform decision-making; and (iii) the stock and flow of timely information are irregular and unreliable. The Sourcebook is specifically targeted towards countries where conditions are less-than-ideal, particularly with respect to the availability of relevant information.

SYNTHESIS

The challenge of understanding reality on the basis of partial information is a recurring theme in the Sourcebook. It is particularly challenging in countries where conditions are less than ideal, that is, where the ability to collect and process statistical data is limited. The Sourcebook cautions against relying on a single source of information and encourages the use of the triangulation process – i.e. combining several sources of information to pick out the key elements of the story. In keeping with the theme of supporting M&E in less-than-ideal conditions, the focus throughout is on assembling recommendations that are pragmatic and practical, rather than abstract and academic. The Sourcebook emphasizes the need to keep things simple and suggests, for instance, that when countries assess their data

needs, they should focus on a minimum set of priority core indicators, rather than on a desired set. It looks at how indicators might be provided and used in conditions where data are limited and capacity to generate them is weak – a situation common to many countries. While the focus is primarily on the monitoring and evaluation of programmes in the agriculture and rural development (ARD) sectors, the guidelines are also relevant to other sectors. Indeed, the approach advocated in the document – which is strongly rooted in the idea of monitoring service delivery and measuring early outcomes – can be generally applied to almost all sectors, and provides an ideal basis for the monitoring of Poverty Reduction Strategies (PRSs) or other national development initiatives.

The Sourcebook reviews best M&E practices under three broad headings: the analytical framework, the data framework and the institutional framework.

ANALYTICAL FRAMEWORK

The analytical framework examines how one measures the impact of the development initiative. What indicators are needed and how are they selected? A complete M&E system must identify and monitor indicators at each of four levels – input, output, outcome and impact. Nowadays, most projects/programmes have a Management Information System (MIS) for tracking inputs and outputs (performance). A fundamental and essential output of the M&E system at this level should be the production of regular performance monitoring reports serving as an input into the preparation of annual work plans and budgets. Tools and approaches such as public expenditure tracking surveys are described in the Sourcebook.

Once systems are in place to monitor performance, attention can turn to the monitoring of results (outcomes and impact) – and this is the area on which the Sourcebook concentrates most. The shift in emphasis from performance to results has profound implications for M&E. Unlike performance monitoring, where data are relatively easily available from internal institutional information systems, measuring results involves turning to the targeted beneficiaries (clients) for information on the project and how it has affected them.

Changes in yield and production levels, whether for crops, fisheries, livestock or livestock products, inevitably feature among the main indicators used for monitoring project outcomes. The Sourcebook suggests that where objective measures are difficult to obtain at the early stage of interventions, farmers' own assessments can serve as useful proxies.

The Sourcebook also shows how a service delivery approach can be used to select indicators which can generate useful, easy-to-measure early outcome measures. It suggests that greater use be made of qualitative indicators, such as access, use and satisfaction.

Finally, there is the question of evaluation. This can be a seriously data-hungry exercise, but for countries with limited capacity, there are ways of getting around the problem. Not all projects/programmes need full-scale impact evaluations, and

where required, they may be carried out without collecting much additional data beyond what has been routinely collected for monitoring purposes – provided the evaluation is carefully planned in advance. Good evaluation will almost certainly involve combining data from various different sources and coming to a considered view on the impact of a particular intervention based on a triangulation process and weighing up of messages – often apparently inconsistent – from different sources.

Nevertheless, for most evaluations and broader planning purposes, the Sourcebook emphasizes the need for a set of basic agricultural and rural sector statistics that extends beyond the service delivery measures. These include basic sector statistics, such as area production and yield data, prices, agricultural input use, public spending on agriculture, the contribution made to GDP by agriculture and GDP per capita. In countries where these are not available, they should be put on a priority list for inclusion in any statistical capacity-building programme. An extended menu of indicators is supplied in Annex 1, which countries can use to help them prioritize and select the most useful indicators for their particular needs. The list is not exhaustive nor is it expected that all countries should adopt and use all the indicators, but it offers a choice and includes examples of good practices taken from different countries around the world.

The discussion of the analytical framework concludes with reference to monitoring and evaluation at the international level. It identifies a set of 19 priority indicators already included in the menu of indicators as core indicators for tracking ARD sector outcomes at the international level. These 19 indicators have been selected on grounds of comparability, availability and relevance. They represent a universal minimum core set and, as far as possible, should be included in all national M&E programmes. Without this minimal commitment at the country level, it is not possible to improve the quality of M&E at the international level, which is one of the purposes of the Sourcebook. But this should not be too onerous a burden, since the same indicators are used to monitor not only at the international level, but also at the national level.

DATA FRAMEWORK

In order to meet the needs of monitoring at each of the four levels (inputs, outputs, outcomes and impact), the M&E system needs to draw on information coming from a variety of different sources. It is not just that each level requires different indicators, but also that the requirements of the users in terms of periodicity, coverage and accuracy vary according to the level of indicator. Input indicators are required to inform short-term decision-making. They therefore need to be produced frequently and regularly – possibly once every 1-6 months. The same applies to output indicators, but here the reporting period can likely be longer. As one moves further up the results chain and starts to collect more information about clients rather than the servicing institution, the task of data collection becomes more complicated. Time must be allowed for clients to become aware of and start using public services. One may see little evidence of outcomes

for the first few years. Therefore, it may be acceptable to build a programme around a reporting schedule of, for instance, 1-2 years. But it is important that the process is initiated at the very beginning of the project with a view to using the first report for establishing the baseline situation. The evaluation of the eventual impact comes much further down the line – often years after the project has been completed. Although the time frame may be more relaxed, the analytical challenge is not, and from the data collection perspective, experience teaches us that it is vital that the outline on how the project is to be evaluated is agreed from the very beginning, since it may involve setting up an experimental design to try to isolate the “with/without” project effect.

The Sourcebook devotes considerable attention to the need for a strong statistical infrastructure and reviews the range of different statistical instruments available.

The most popular and obvious instrument for monitoring outcomes of ARD programmes is the household survey. It provides data that can be disaggregated to show results for different population groups and has the advantage of providing information on both beneficiaries and non-beneficiaries. There are a number of different household survey models that can be used, each with its own strengths and weaknesses. The Sourcebook assesses their relative strengths and weaknesses and approximate costs. The most complete coverage is provided by the population census. Although obviously not appropriate for day-to-day monitoring, the census is important because it provides the framework for almost all other household survey activities, including agricultural censuses and surveys. The latter are extremely relevant to the monitoring of ARD programmes because they are usually the only means of monitoring changes in crop production levels and yields. Integrated multi-topic household surveys are another form of enquiry that has become increasingly popular. They are particularly good as baseline surveys that can be used to measure poverty levels, identify potential problems in need of attention, and generally understand the way in which households establish mechanisms to cope with difficult living conditions. The big disadvantage is that they are difficult surveys to undertake, and many countries have neither the analytical nor the survey capacity to successfully carry out such large-scale complex surveys on a regular basis. Lighter and more rapid household surveys are, however, becoming increasingly popular. Service delivery surveys have been used in market research for a long time, but are relatively recent additions to a National Statistical Office (NSO)’s repertoire of surveys. They are extremely well-suited to monitoring early results. They are also easy to implement and can be repeated annually without disturbing any other survey work that the NSO may be undertaking.

In addition to household surveys, a good M&E system will use a wide range of other tools. These can include community surveys, which may be conducted both on probability and non-probability samples, and qualitative surveys and studies, including participative assessments, focus group discussions and rapid appraisals such as windscreen surveys. Institution-based surveys, such as Quantitative Service

Delivery Surveys (QSDSs), can also play an important role in highlighting supply-side constraints, as can the analysis of administrative records.

The main message to emerge from the Sourcebook is that no single instrument can meet all needs and that any monitoring system will most likely acquire indicators from several sources – both formal and informal. Since it can take a while for the necessary capacity to be built, the Sourcebook offers a number of possible shortcuts for countries with less developed statistics systems.

In many countries, NSOs have found themselves caught in a vicious circle in which users have become disillusioned because the statistical products are late, inaccurate and filled with blanks. In a number of cases, this has led users to become dismissive of the efforts of the NSO, and in the process, to stop providing feedback on how databases could be improved. The inevitable knock-on effect is that resources for statistics are reduced and, as a result, so are NSO capacities. However, the future looks more promising and the signs are that with some assistance, NSOs will be able to rebuild capacity and meet the new information demands required by the monitoring of national development strategies.

INSTITUTIONAL FRAMEWORK

The final challenge in building up M&E competences is neither technical nor conceptual, but managerial. It concerns ensuring that the required incentive structure and institutional capacity are created to be able to perform this work. Whether countries already have an active ongoing M&E programme or whether they are starting from scratch, they need to regularly review all ongoing M&E activities. This may unearth a number of apparently duplicating and conflicting structures, but the goal should be one of inclusion not exclusion, and of creating a network of institutions engaged in M&E.

At the core, there needs to be a central M&E unit with the authority to coordinate the different initiatives. One of the more important functions of the unit should be to promote and encourage the demand for M&E. At the same time, it needs to help establish stronger links with data suppliers within the National Statistical System (NSS).

Despite the numerous areas of common interest, in many countries there appear to be two distinct and separate communities of practice – the M&E community and the statistics community. Both may be working on parallel issues but not necessarily communicating or working together. At the same time as the growth of interest in the M&E of national development programmes, there has been a similar interest in the rehabilitation of NSSs. The NSS comprises all the institutions and agencies that contribute in some way to the bank of national statistical data, which includes line ministries, Customs and Excise and the Central Bank, among others. The apex institution for the NSS is the NSO. Many countries are now developing National Statistical Development Strategies (NSDS) in such a way that they are integrated into national development policy processes. This ties in closely with the ideas underpinning the development of national M&E capacity.

THE ROLE OF DEVELOPMENT PARTNERS

Donors have been among the strongest advocates for establishing good M&E procedures and for building up M&E capabilities. They have also provided strong support to the strengthening of national statistical capacity, but in many cases, their efforts have been counter-productive as a result of a failure in coordination. However, all major donors have now subscribed to the Marrakesh Action Plan for Statistics (OECD, 2004), in which donors commit themselves to working collaboratively to support countries in the preparation of NSDS.

EMERGING ISSUES

One cannot leave the discussion of the evolving role of M&E without making reference to three new and growing challenges. The first is the impact of devolution and decentralization on M&E. Many countries now pursue broad decentralization policies aimed at bringing the government closer to the people and enhancing transparency and accountability. This has profound consequences for M&E, which is now obliged to provide indicators at a much lower level of disaggregation. When the data source is administrative records, this may not present much of a problem. But when the source is a statistical survey, it can require dramatic increases in sample sizes, which may call for a major rethinking of how data are to be collected. The second challenge concerns the involvement of communities themselves in M&E. As interest in community-driven development projects continues to grow, so too does the demand for community-driven M&E in which the communities themselves take charge of their own M&E. This is likely to be an area in which major methodological developments will occur. Finally, there is the challenge of the monitoring and evaluation of ARD programmes at the global or international level. Monitoring international/global goals is the responsibility of the international development institutions, including the specialized agencies of the United Nations, the World Bank and the International Monetary Fund (IMF), but ultimately these entities depend on the NSSs to provide the basic data. The relationship between national and international institutions engaged in monitoring is not hierarchical, but rather, complex and symbiotic. Ultimately, the global M&E network is only as strong as its weakest link. International agencies therefore have a vested interest in seeing that the capacity of national institutions is strengthened.

SETTING UP AN M&E STRATEGY IN AGRICULTURE AND RURAL DEVELOPMENT

The Sourcebook makes the point that a fully evolved M&E system is more than a simple tracking system to measure performance and outcomes. These activities need to be put into the context of a cyclical approach to management in which:

- planning involves the articulation of strategic choices in light of past performance;

- implementation includes ongoing performance monitoring and periodic evaluation that provide opportunities for learning and adjustment;
- reporting on results is used both for internal management and for external accountability to stakeholders, including civil society. The reporting phase also provides managers and stakeholders with the opportunity to reflect on what has and what has not worked – a process of learning and adjusting that feeds into the next planning cycle.

The Sourcebook, in its final chapter, describes the key elements of an ARD M&E strategy and sets out the key steps that need to be followed to set it up, namely:

- Assessment of current M&E capacity and diagnosis.
- Review of indicators using the methodology described in Chapter 2 and, where appropriate, the suggested indicators provided in Annex 1.
- Review of current data, sources and gaps. The assessment should include a review of the quality and timeliness of the data and should draw on information contained in Chapter 3.
- Develop action plans linking together the M&E activities of all the institutions involved – as described in Chapter 4.
- Review resource requirements.
- Define a system to monitor the performance of the M&E action plan.

What is, in effect, being proposed in the Sourcebook is that countries should define a strategy for developing national M&E capacity as part of their overall ARD strategy. This would result in a better understanding of what works and what does not, which will lead directly to better planning of future programmes and projects. It will also lead to better programme implementation by providing timely warnings suggesting how resources may need to be reallocated when actual results are deviating from expected results.