

CHAPTER 1

THE EVOLUTION OF M&E IN DEVELOPMENT

The chapter opens with the question “What is M&E?” and then demonstrates how M&E has different meanings for different groups. The chapter then describes how M&E has evolved over the last 20 years from its early beginnings as a project-based evaluation tool to its current form, which is used for tracking multisectoral national development programmes such as Poverty Reduction Strategies.

WHAT IS M&E?

In the old story of the blind men and the elephant, a group of blind men touch an elephant to determine its true nature. Each one touches a different part. The one who feels a leg says the elephant is like a pillar; the one who feels the tail says the elephant is like a rope; the one who feels the trunk says the elephant is like a tree branch; the one who feels the ear says the elephant is like a fan; the one who feels the belly says the elephant is like a wall; and the one who feels the tusk says the elephant is like a solid pipe. They each claim to know what an elephant is but they are in complete disagreement. All are partially right, yet all are wrong.

The story of the blind men and the elephant could apply to M&E. Ask six people what M&E is and you get six different answers! It means different things to different people: M&E is a management tool; M&E improves planning; M&E is applied research; M&E is a tool to improve governance and accountability; it empowers communities; it monitors global goals. In fact, it covers all of the above and includes project supervision, financial monitoring, surveys and statistics, MISs, social analysis, and the setting and tracking of national development goals. Yet, it is more than the sum of its component parts.

The story of the blind men and the elephant is also relevant to M&E in another way. It illustrates how difficult it can be to understand reality on the basis of partial information. This underlines one of the key messages of the Sourcebook, which is to emphasize throughout the importance of sharing and triangulating information from different sources, and to be wary of relying on a single source of information. This applies equally to qualitative and quantitative information.

Different sources have their own individual strengths and weaknesses. In the area of poverty monitoring, for instance, the messages derived from qualitative studies based on participant observation often yield results that are seemingly at odds with the findings from “objective” statistical household surveys. The temptation is to reject one (usually the qualitative data) as being wrong. This would probably

Box 1. Definitions of monitoring and evaluation: 1984-2002

1984

Monitoring is a continuous assessment both of the functioning of the project activities in the context of implementation schedules and of the use of project inputs by targeted populations in the context of design expectations. It is an internal project activity, an essential part of good management practice, and therefore an integral part of day-to-day management.

Evaluation is a periodic assessment of the relevance, performance, efficiency and impact of the project in the context of its stated objectives. It usually involves comparisons requiring information from outside the project – in time, area or population.

IFAD, 2002

2002

Monitoring can be defined as “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”. Thus, monitoring embodies the regular tracking of inputs, activities, outputs, outcomes and impacts of development activities at the project, programme, sector and national levels. This includes the monitoring of a country’s progress against the Millennium Development Goals (MDGs) or other national measures of development success.

Evaluation can be defined as “the process of determining the worth or significance of a development activity, policy or program to determine the relevance of objectives, the efficacy of design and implementation, the efficiency or resource use, and the sustainability of results. An evaluation should (enable) the incorporation of lessons learned into the decision-making process of both partner and donor.”

OECD, 2002

be a mistake. The measurement and monitoring of living standards is a highly complex undertaking because of the multifaceted nature of the subject matter. When trying to interpret messages coming from different sources, it may at times seem as if one is trying to compare apples and oranges. Closer inspection and comparison of the two sources, however, often reveal important insights and show that far from contradicting each other, they actually highlight different aspects of poverty and provide complementary information. The key point is not to misuse any one instrument and expect it to answer questions that it was never designed to answer.

The first task of the Sourcebook, therefore, is to ensure that everyone has a common understanding of the issues that M&E can legitimately be expected to address. Various texts have defined M&E differently, which leads to more confusion. Among the earlier attempts, the clearest and least ambiguous definitions were found in the *Guiding principles of the design and use of monitoring and evaluation in rural development projects and programmes*, produced by IFAD in 1985 in cooperation with FAO and the World Bank (IFAD/FAO/WB, 1985). Box 1 compares the definitions established in 1984 with those revised and updated by the DAC Network on Development Evaluation (OECD, 2002) almost 20 years later.

The language is different but the concepts are broadly similar. What has changed, however, is the way in which the M&E concepts are applied. In the early days, the focus was on the project – a relatively well-contained development initiative with a limited time frame and clearly articulated goals. Today, however, the focus of M&E efforts is much broader and encompasses the M&E of sectoral plans and programmes, national development strategies, and, indeed, the international Millennium Development Goals.

Another important point to note is that, in both the earlier and the later definitions, the idea of M&E as an audit-like surveillance tool is excluded. Where there is an M&E unit, rather than being treated as an external agent, it is integrated into the project management structure and serves as a resource for supplying key information on project implementation and delivery. The function of the M&E unit is seen as assisting management by establishing and maintaining appropriate MISs and ensuring that they produce reliable data in a timely manner. Good management requires a good MIS and that the monitoring function is carried out using the data from within the MIS. Such a system includes the basic physical and financial records, the details of inputs and services provided to the beneficiaries or clients (for example, credit and extension advice) and data obtained from surveys and other recording mechanisms designed specifically to collect information from the service users.

M&E has evolved from being a set of project management tools to becoming a core element of national strategies for reducing poverty.

Monitoring and evaluation are closely linked but separate activities.

Evaluation is seen as a separate function but linked to monitoring. Evaluations can be simple or complex. There are several different kinds of evaluations, ranging from short desk reviews of documents and performance audits, to full-scale impact evaluation. Impact evaluation has a critical role to play in increasing knowledge about what works and what does not. Impact evaluations can be immensely valuable but are not easy to carry out. They draw on the MIS to provide data for making comparisons over time and against comparable “control” information, but they also require information from the clients – the intended beneficiaries. This requires baseline information. In the beginning, it was implicitly assumed that the project M&E units would undertake baseline surveys of their own with the understanding that the survey would be repeated at the end of the project and any differences would then be attributable to the project itself. In most cases, this proved to be much more difficult than anticipated. In many cases, the survey was overambitious and took years rather than months to complete. At times, the second survey was never undertaken, or if it was, the size of the combined sampling and non-sampling errors was found to be larger than the real change that the surveys were meant to detect.

Even today, the relationship between monitoring and evaluation continues to be the subject of discussion. At one end of the spectrum, there are those who put the primary focus on monitoring, and see M&E principally as a management support system whose main concern is to ensure the timely production of appropriate indicators. At the other end of the spectrum, there are those with an equally strong argument that the primary function should be to carry out effective impact evaluation from which lessons could be learned for the future. Then there are those who feel that M&E systems should be capable of doing both. This middle path is the one that is usually taken – a sensible compromise where one must, however, be continually aware of the risk of spreading resources too thinly in trying to achieve multiple objectives and ultimately satisfying none.

The measure of a good M&E system is customer satisfaction.

To summarize, the basic principle is that monitoring is an ongoing activity and evaluation is periodic, carried out at specific times during the project cycle (annual, mid-term, terminal) or indeed after the completion of the project (impact evaluation). In broad terms, M&E are activities whose primary function is to provide appropriate information at the right time to users with decisions to make and to improve their decision-making as a result. M&E, like all other services, can only function effectively if there is a demand for

it. How can one know whether the system is working correctly or not? In the long term, one would seek evidence of better planning, resource allocation and administration of development programmes as a result of learning from experience. In the short run, the answer is satisfied users. If there is a growing number of people who are aware of M&E data and also a growing number of people actually using the data, then one may infer that the system is providing a useful service.

WHO ARE THE USERS?

The more open or inclusive the system of government, the broader the range of users is likely to be. At the start, the focus of the M&E reporting system may be on budget management and performance budgeting, but as the programme or project grows and the number of beneficiaries increases, so does interest in the M&E data. Users include those who have a financial or management interest in the project (donors, government), as well as the beneficiaries, the media, civil society at large and their representatives (parliament).

At the beginning, however, it can be hard to raise any interest at all. In the early days, in many countries, the demand for good M&E information originated entirely from outside sources. The donors were driven by an electorate at home that needed to be satisfied that aid funds were being used for the intended purposes and were achieving results. In the developing countries in which the M&E systems were being installed, however, there was generally little interest. Even in projects that included a donor-driven M&E component, managers were ambivalent about its value and tended to see M&E units as a drain on their resources, or even worse, as an informant imposed from outside. We have moved a long way since then, but still, without in-country demand, no system can be sustainable. Therefore, one of the first requirements for successful M&E is to nurture and cultivate the demand. This is likely to mean taking measures to initiate a strong advocacy programme to inform potential user groups about the benefits of a results-driven environment. Consequently, M&E has become an important pillar of the PRS and not just a marginal activity. As shown in Box 2, the PRS can underline the need for good M&E data to: (i) support budget decision-making; (ii) help with policy formulation and programme development; (iii) support the management of sectoral programmes; and (iv) signal whether the programmes are genuinely contributing to an improvement of living standards and well-being in the country. However, the process of reorienting a country or culture to value a results-oriented government system can be a long and arduous process. In summary, monitoring information and evaluation findings can contribute to sound governance in a number of ways, but primarily through evidence-based policy-making (including budget decision-making), policy development, management and accountability.

Box 2. How M&E findings help governments and stakeholders?

M&E findings:

- support policy-making, especially budget decision-making, performance budgeting and national planning. These processes focus on government priorities among competing demands from citizens and groups in society. M&E information can support the government's deliberations by providing evidence of the most cost-effective types of government activity, such as different types of employment programmes, health interventions, or conditional cash transfer payments. Terms that describe the use of M&E information in this manner include evidence-based policy-making, results-based budgeting and performance-informed budgeting;
- help government ministries in their policy development and policy analysis work, and in programme development.;
- help government ministries and agencies manage activities at the sector, programme and project levels. This includes government service delivery and staff management. M&E identifies the most efficient use of available resources and can be used, for example, to identify implementation difficulties. Performance indicators can be used to make cost and performance comparisons – performance benchmarking – among different administrative units, regions and districts. Comparisons can also be made over time that help identify good, bad and promising practices, which can prompt a search for the reasons for this performance. Evaluations or reviews are used to identify these reasons. This is the learning function of M&E and is often termed “results-based” or “results-oriented management”;
- enhance transparency and support accountability relationships by revealing the extent to which the government has attained its desired objectives. M&E provides the essential evidence necessary to underpin strong accountability relationships, such as the government to the Parliament or Congress, civil society and donors. M&E also supports the accountability relationships within government, such as between sector ministries and central ministries, among agencies and sector ministries, and among ministers, managers and staff. Strong accountability, in turn, can provide the incentives necessary to improve performance.

World Bank, 2007

How M&E HAS EVOLVED

At this stage, a historical learning exercise may be useful. In the following description of how M&E has evolved over recent decades, this process has been grouped into several distinct phases for the purpose of clarity. This is an oversimplification and disguises the fact that progress is neither sequential nor linear, but it does help to show how ideas have evolved and how expectations have expanded over the years.

In the beginning: project-based M&E

The first signs of interest in M&E for ARD projects became evident in the mid-1970s. At that time, interest was strictly project-based and there was general agreement that projects could be better designed and managed with a strong M&E programme. In many cases, this involved the establishment of a dedicated M&E unit.

During this early evolutionary phase of M&E, its main purpose was to serve as a management tool that would provide timely feedback and give warning whether the project was on track or not. While paying lip service to the need for measuring outcomes, the focus of interest was on the monitoring of inputs and outputs. The project document was treated more like a “blueprint” than a “roadmap”. If the planning had been correctly done, then the main purpose of M&E was to provide timely feedback that the project was being implemented in line with expectations – and if not, to send a quick warning. This is still an important aspect of M&E even today.

In the 1970s, interest in M&E was strictly project-based: its main purpose was to serve as a management tool.

Expanding horizons: programme and sectoral M&E

By the early 1990s, a change was taking place in how development aid was being administered, leading to a shift in focus from the project to the sector-wide programme. Programmatic aid, whether in the form of loans or grants, was becoming increasingly common, since it was seen that project-based assistance was failing to deal with the larger systemic problems and was not creating an effective investment-friendly environment necessary for sustainable development and long-term raising of living standards. The effect was not so much that projects were discontinued – indeed they continued to thrive – but that a sector-wide approach (SWAP) became increasingly popular as a means of promoting and coordinating sector-wide and national development planning. These development models potentially gave more flexibility to governments and programme executing agencies, but good reporting and feedback systems had to be conceived of as an integral part of the programmes.

The expansion from project to programme-level support had enormous implications for the M&E system.

One of the results of this SWAP was the recentering of many M&E activities from the project level to the sectoral level. Monitoring and evaluation became functions of sectoral ministries and appropriate M&E units were established at the ministry level. Sometimes, the sectoral units entirely replaced the project units; sometimes they did not. A network of M&E units were created, in which project units either copied their reports or sent them directly to the sectoral M&E unit. The

The focus turns to the beneficiaries, which requires better data and more tools.

nature of the relationship between the project units and the sector unit varied substantially from country to country. In some, it was rigid and hierarchical; in others, the relationship was much looser. But the old custom of allowing each project to design and develop its own M&E procedures was in general replaced with a more centralized approach that would ensure that all programmes and projects followed the same procedures and reporting formats so that statistics could be compiled into sector-wide

reports. Development partners also had to be prepared to accept a standard format rather than insist that their own individual reporting formats be used.

In the 1990s, the idea of results-based management was also becoming popular. The consequence was a shift in emphasis away from the monitoring of inputs and outputs to the measurement of “results” – a much more difficult task. This expansion of expectations was a significant change from before.

The early involvement of NSOs was not particularly successful.

Up to that time, it was possible for much of the data to be generated from internal reporting systems. Then, in order to measure the results of project activities, the focus of M&E had to switch from the project to the client or intended beneficiary. It thus became necessary to call on a much wider range of data tools and sources. Surveys and beneficiary interviews in particular would need to be undertaken, which required a level of expertise and training not generally

available in project M&E units, or even in the M&E units of sector ministries. For the most part, M&E staff did not have the time, training or the resources to tackle this kind of work. The involvement of new players with more technical expertise was needed.

One new player was the NSO. The primary function of an NSO had always been to act as the ultimate source and repository of all official national statistics. In most countries, they were established as a government body with only limited autonomy. Their most important outputs were national accounts, an annual statistical abstract and the published results of whatever survey or census they happened to have undertaken recently. In many countries, it seemed to be the only institution with the knowledge and capacity to collect and process data on the scale needed by the project. It was thought that either it would be possible for projects to “piggyback” onto the NSO’s household survey infrastructure

and to use the NSO survey as a means of measuring project results, or it could undertake special surveys specifically for the project. In both cases, the outcome was generally disappointing. Statistics offices were, on the whole, overextended and under-resourced, and failed to rise to the challenge. Adherence to timeliness and respect for deadlines were not qualities commonly associated with under-resourced NSOs – nor was adaptability. Another problem was that the data supplied were generally too “macro” and not sufficiently disaggregated for M&E purposes. While their data could make a contribution to the overall performance of national and sectoral development programmes, they were generally not specific enough to be helpful in measuring the outcome of specific development interventions. Either they were not repeated with sufficient regularity to allow for comparisons over time, or they could not be sufficiently disaggregated to allow for comparison between different subgroups of the population. The dialogue between the national data provider and the data user was not easy, and led to frequent disappointments.

The arrival of poverty monitoring

Another force that started to emerge during the mid-1990s was concern about the issue of poverty. While the primary goal for a developing country had traditionally been “development through growth”, it changed in the 1990s to “growth and poverty reduction”; it was not enough to aim for wealth alone. It now became increasingly clear that this had to include a fight against poverty and protection for the poorest. A new branch of monitoring activity was required: poverty monitoring. This was a complex and challenging undertaking that, for the most part, was built around the tracking of living standards with a view to anticipating the direction in which they would likely move as a result of macroeconomic policy. One of the underlying driving forces was the concern that structural adjustment programmes – introduced in many countries in order to redress economic imbalances and improve international competitiveness – could be imposing undue hardship on some of the most vulnerable elements of the population.

In the 1990s poverty monitoring was introduced to study the effect of economic development on living standards.

Most countries started out with very little knowledge or capacity to monitor poverty. Support from donors focused primarily on assistance in the design and implementation of multi-topic household surveys, which included the measurement of household consumption as the indicator of choice for measuring poverty. Only NSOs had the capacity to undertake such large-scale national household surveys, but even then, in most cases, they did not have the capacity to analyse them.

As time progressed, qualitative and quantitative tools were added, including participatory poverty assessments, poverty mapping and the tracking of core

indicators over time. These involved bringing on board other institutions, including academic institutions and NGOs. To coordinate all these activities, countries started to establish National Poverty Monitoring Units. The results of these efforts were mixed, but overall capacity was being built. What is interesting, however, is that the building up of a national poverty-monitoring capacity was kept distinct and separate from other M&E capacity-building efforts, and there was very little communication between them ... until the new millennium.

Monitoring Poverty Reduction Strategies: building national M&E capacity

By the turn of the millennium, poverty alleviation had moved from being a marginal issue to being a central concern for almost all countries. A target of

M&E becomes a key agent of development in its own right.

halving global poverty by 2015 was enshrined as the first Millennium Development Goal. At the country level, the National Poverty Reduction Strategy (PRS) was introduced to serve as a framework for promoting the vision of “pro-poor growth” (Box 3). The earlier experiences of setting up country-level poverty monitoring systems were to prove critically important for the introduction and successful implementation of national PRSs. The poverty assessments provided the means of identifying

where the most vulnerable were located.

The new millennium saw the bringing together of project- and sector-based M&E efforts with poverty monitoring activities. The result was the emergence of national M&E programmes centered around the monitoring of PRS results. At this stage, M&E started to emerge as a key agent of development in its own

Coordinating M&E activities across and within sectors remains a challenge.

right, and an essential component of the PRS. In-country demand, which had previously been limited, started to expand – and with it, recognition emerged that M&E information should be not just a tool for policy-makers and planners, but should be made readily available to members of the public and to civil society. In this way, the M&E system started to become a tool for promoting good governance and accountability.

Box 3. Poverty Reduction Strategy Papers (PRSPs)

Poverty Reduction Strategy Papers (PRSPs) are prepared by governments in low-income countries through a participatory process involving domestic stakeholders and external development partners, including the International Monetary Fund (IMF) and the World Bank. A PRSP describes the macroeconomic, structural and social policies and programmes that a country will pursue over several years to promote broad-based growth and reduce poverty, as well as external financing needs and the associated sources of financing.

What is the purpose of PRSPs?

The world economy has grown steadily in recent decades, bringing widespread prosperity and lifting many millions out of poverty, especially in Asia. Nevertheless, in the next 25 years, the world's population is projected to grow by about two billion people, most of whom will be born in developing and emerging market economies. Without concerted efforts by countries to help themselves through sound policies and by the development community to increase its support of the countries' own efforts, many of these people will be doomed to poverty.

The PRSP approach, initiated by the IMF and the World Bank in 1999, results in a comprehensive country-based strategy for poverty reduction. It aims to provide the crucial link between national public actions, donor support and the development outcomes needed to meet the United Nations' Millennium Development Goals (MDGs), which are aimed at halving poverty between 1990 and 2015. PRSPs provide the operational basis for Fund and Bank concessional lending and debt relief under the Heavily Indebted Poor Countries (HIPC) Initiative. They are made available on the Web sites of the IMF and World Bank by agreement with the member country.

Core principles of the PRSP approach

Five core principles underlie the PRSP approach. Poverty reduction strategies should be:

- **country-driven**, promoting national ownership of strategies through broad-based participation of civil society;
- **result-oriented** and focused on outcomes that will benefit the poor;
- **comprehensive** in recognizing the multi-dimensional nature of poverty;
- **partnership-oriented**, involving coordinated participation of development partners; (government, domestic stakeholders, and external donors);
- based on a **long-term perspective** for poverty reduction.

IMF Factsheet, September 2005

