

Title: Strengthening Capacity to Monitor and Evaluate Agricultural and Rural Development Programs

Author: Marchant, Timothy

World Bank/FAO
Clock House, Berwick St James
Salisbury, SP3 4TN, UK
tjmarchant@btinternet.com

Abstract

Building the capacity to measure results and to use that knowledge to learn what works and what doesn't, or how to make things work better is not an easy task, and requires the existence of a strong monitoring and evaluation (M&E) capacity backed up by an effective statistical system capable of generating timely and relevant development indicators

FAO and the World Bank have joined together to support the preparation of a sourcebook for countries needing help with establishing or strengthening monitoring and evaluation capacity. A draft document, Selecting Monitorable Indicators for Agricultural and Rural Development Programs - Measuring Results in less-than-ideal Conditions, was prepared in 2006 and the methodology and the indicators were validated in five countries. A final version of the Sourcebook, incorporating the lessons learned from the country experiences, is due for dissemination by the end of this year.

The sourcebook is intended to assist countries to strengthen their capacity to monitor agricultural and rural development programs and to measure results. It should serve as a Handbook for development practitioners working in Agricultural and Rural Development (ARD) and provide them with certain basic tools to help with the selection and use of core indicators for monitoring and evaluating projects and programs. It focuses on particularly the measurement of results (outcomes and impact). At this level indicators can be quite hard to define since the underlying data are often not available. The Sourcebook suggests ways or simplifying the process for instance by using, where possible, a "service delivery" approach.

The sourcebook also addresses supply-side issues and the need for building up capacity of the national statistical systems. Various statistical instruments are reviewed, and a menu of indicators that can be used for monitoring and evaluating agricultural and rural development programs is also included.

Measuring results in less-than-ideal conditions

The universal acceptance of the millennium development goals (MDGs) represents a global commitment to lift the poorest of the poor out of poverty. It establishes a demand for M&E at the very highest level. Not only will it be necessary in 2015 to report on whether or not the goals have been achieved, but it is also important that, well before then, the mechanisms are in place to track progress towards their achievement and to alert stakeholders to issues of concern where countries or regions are clearly off-track. To achieve the millennium development goals the international community must assist more than one billion people out of extreme poverty, of which 70% live in rural areas and depend on agriculture for their livelihood. The challenge is to understand how, where and when agriculture can make the greatest contribution to achieve the MDGs. The Global Donor Platform for Rural Development is an association of development partners committed to supporting and strengthening

agricultural and rural development projects and programmes around the world. It has decided that one way of doing this is to support the strengthening of monitoring and evaluation of agricultural and rural development projects and programmes, and consequently it is financing the production and dissemination of a new M&E sourcebook. How will this publication be any different from any other technical book on M&E? The answer lies in the subtitle "Measuring results in less-than-ideal conditions". The sourcebook takes as its starting point the fact that in many countries monitoring and evaluation capacity (and, in particular, statistical capacity) is extremely weak. For such countries it is simply impractical to propose elegant, top-of-the-line, M&E programs. What they need is a minimalist system that can supply just enough information at the right time to check whether their agricultural and rural development programs are on track or not -- not one that imposes totally unreasonable demands on the national statistical system - and on the usually under-resourced and overstretched national statistics office (NSO). The sourcebook aims to respond to that need.

In this paper I shall be selective when describing the contents of the sourcebook and shall pick a few examples of ways in which the sourcebook suggests for simplifying M&E activities yet at the same time make them more relevant to user needs.

The Sourcebook starts with an overview of M&E and describes how it has evolved in the development context. It describes how M&E has grown from being essentially a project-based management tool to a country-based program for monitoring the outcomes of national development strategies. It describes the spectrum of views between those who saw M&E as a management tool and those who saw it more as a research tool. It shows how the growth in demand for poverty monitoring has led to a renewed focus on building up national statistical capacity.

The sourcebook goes through a number of issues related to the monitoring and evaluation at agricultural and rural development programs, and groups in under three headings: the analytical framework, the data framework, and the institutional framework.

The **Analytical Framework** covers what are essentially the demand-side issues: who needs M&E data, and what will they do with them. The sourcebook shows how the use of the logical framework approach (log frame) can help with the selection of monitoring indicators and with Fleur the prioritisation of information needs. It describes how the original emphasis was on tracking the input and output indicators (performance monitoring) but how attention then turned to monitoring results: outcomes and impact. At this level projects and programmes are experiencing difficulties particularly in the less well-off countries. What can they do about it?. The Sourcebook suggests that in many cases the gap can be filled by using a simple service delivery approach to selecting indicators. A service delivery approach considers that most projects have one thing in common. They are essentially vehicles for making a product or products available to a target population. The concept of the "product" is a broad one. It can be:

- a tangible product such as a loan, a rural road, or a package of technological innovations for increasing yields;
- a service, such as an extension program, local health care, or land registry service;
- something more abstract, such as “an enabling environment” or a “community development project”.

What a service delivery approach then does is to focus on the clients -- the potential users of the product -and find out from them their views on the ‘product’ and the extent to which they make use of it. The basic questions that need answering are:

- Do clients have **access** to this product? (Do they know about it? Is it physically accessible to them? Can they afford it?)
- Do they **use** this product?
- If yes, are they **satisfied** with the quality of the product?
- If no, why not?

From these questions it is then possible to generate three basic indicators: access, use and satisfaction. Though these are quite simple indicators they have a number of qualities which make them eminently attractive as outcome indicators.

- They are relatively quick to process..
- They can be collected regularly, say once a year
- They can be disaggregated so that comparisons can be made between different groups
- they are easy to collect

The query may be raised about the validity of using “satisfaction” as a measure of success. Can one really trust the respondent to give an honest answer? How can one quantify such a subjective notion? The Sourcebook argues that there is no reason why a subjective assessment such as satisfaction, cannot be a valid indicator to include as one of the early measures of outcomes. In fact who better to evaluate a product than the user him or herself. Monitoring and evaluation are not exact sciences but involve a process of triangulation – of picking up information from various sources and of combining and comparing them to arrive at the most probable assessment. The respondent’s opinion is as valid as any other source of information, and even though it is subjective, there are standard procedures whereby it can be quantified.

The sourcebook differs from many other publications on monitoring and evaluation in that it looks at the issues very much from the point of view of the data provider and the question is consistently asked “how difficult would it be to provide this information?”. In promoting the use of service delivery indicators there is no suggestion that other measures of project outcomes should be dropped. Production and yield indicators will continue to be required for most ARD projects, but they are problematic and long-term and it may take a number of years before lessons can be drawn from them. Access, use and satisfaction indicators provide very useful advance information.

Under the heading of the **data framework** these issues are discussed in more depth. Most agricultural projects share the goal of raising of agricultural output.

One would think therefore that the appropriate indicator for this would be to measure ‘yields’ – calculated as the ratio between production and area cultivated. Unfortunately this is not usually the case and for two reasons. The first is essentially a statistical one and centers around the issue of time series analysis. The problem is that agricultural production fluctuates and can vary significantly from one year to the next – primarily, but not exclusively due to the strong effects of rainfall, or the lack of it. This phenomenon is particularly acute in non-irrigated conditions. As a result it is frequently not possible to detect any change in the trend until a number of years have passed – as many as seven or eight years. But that’s not the only difficulty:, there is also the problem of measurement errors – errors associated with the measurement of smallholder crop areas and crop production. The standard methodology involves gaining the basic measurements by a walking around the perimeter of the plot using a compass and chain (or odometer) to calculate the area, and then laying down one or two randomly placed squares and harvesting the yield from the squares to cut collect production. The procedure is time-consuming and prone to errors. These can be very large particularly in Africa where traditional plots frequently include multiple crops, irregular planting and ill-defined even non-existent plot boundaries. But one of the more encouraging messages highlighted in the sourcebook is a possibility of cleaning reliable yield and production estimates directly from the farmer directly by simply asking him to state how much he produced . Simple, time-saving and cost-effective, and there is strong evidence to suggest that the farmers’ estimates produce significantly better results than so-called objective measurements.

Another issue that has taken up in the sourcebook is the measurement of poverty in countries with limited statistical and analytical capacity. The ultimate goal of nearly all ARD projects and of national poverty reduction strategies is to reduce the level of poverty, i.e. to increase rural incomes as a whole and at the same time to reduce income disparities between the rich and the poor. If the measurement of agricultural production was thought to be difficult, the measurement of living standards is even more challenging. In order to track the very first MDG poverty indicator – “% of the population living on less than one dollar a day” – a detailed household survey is required and enormous volumes of data to be collected. The survey may involve multiple visits to households and the collection and processing of 200 or more items of data from every single sample household to compute an estimate of household consumption. Further information then has to be provided on all household members, including their age and gender, in order to estimate per capita consumption. More data is then needed on comparative prices before the complex analytical task can begin of finding out who is and who is not below the poverty line. In most countries this is not the sort of indicator that can be realistically measured more frequently than once every five years or so. This leads one, once again, to be on the lookout for alternative measures or methods that could be applied in countries where conditions are less than ideal. The key question to be asked is whether it is really necessary to be monitoring changes in poverty levels? There is no doubt that it is required for assessing whether the first MDG target is achieved, but does it help with the monitoring and evaluation of an agricultural project? The sourcebook suggests that, where resources are limited choices have to be made, and that there may be more urgent data demands. It suggests that, instead of trying to answer the question

“has poverty increased or decreased? the more immediate need is to answer the question “are the core ARD services reaching the poor as well as the non-poor?” The first question, though interesting, may not be answerable for a number of years. The second question however can be tracked from year to year relatively easily and the answer can be used to help decide on a better reallocation of public sector resources in order to accelerate the effects of poverty reduction policies. This is an important shift in emphasis in that it focuses attention on the provision of services rather than on the measurement of poverty. But is it any less demanding in terms of data needs than the original question? It still requires the classification of households into the poor and non-poor.

The sourcebook suggests several measures that can be taken to simplify the process. First, in order to use the classic methodology described above, a poverty line, usually based on minimum food and non-food requirements, needs to be established. This is, again, a data-hungry exercise. An alternative and somewhat simpler solution is to use a relative concept of poverty. For instance instead of having a fixed poverty line one could simply rank households from richest and poorest and classify, say, the bottom 20% as being “the poor” and the bottom 10% as the ‘ultra-poor. In one fell swoop all the complexities of establishing the poverty line are removed and the analytical task is simply to compare the services reaching the bottom 20% compared with those reaching the rest of the population.

But the problem remains that in order to obtain the consumption-based measure of poverty (i.e a measure of household consumption) a household survey is required which includes the collection of household expenditure and consumption, and this will need to be updated periodically. For many countries this is simply not practicable. Yet again this may not be an unsolvable problem. A number of countries are now experimenting with much lighter household surveys which do not involve the collection of consumption data but collect specific easy-to-measure indicators of household well-being such as asset ownership, number of literate adults; number of children malnourished; housing quality and mean number of persons per room; adults unemployed etc.. These are used to create a composite poverty index which is then used to rank households according to the index and group them into poverty deciles. Once this point has been reached, comparisons can again be made between poor and non-poor households. More importantly, these same light surveys can also be used for collecting access, use, and satisfaction indicators for a range of different public services. This makes them into a very powerful M&E instruments.

The sourcebook is however generally wary of putting too much faith in a single source of information. It repeatedly reminds the reader of the weaknesses, as well as the strengths of different instruments. One of the key messages of the Sourcebook is to emphasise 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 information and to quantitative. Different sources have their own individual strengths and weaknesses. In the area of poverty monitoring for instance it is often the case that the messages derived from qualitative studies based on participant observation yield results that are seemingly at odds with the findings from

so-called objective statistical household surveys. The temptation is to reject one or the other (usually the qualitative data) as being wrong. This would probably be a mistake. The measurement and monitoring of living standards is a highly complex undertaking because of the multifaceted nature of the subject matter. Closer inspection and comparison of the two sources will often reveal important insights and show that far from contradicting each other, they are in fact highlighting different aspects of poverty and are providing complimentary information. The key point is not to misuse any one instrument and expect it to answer questions that it was never designed to answer.

However it takes more than tools to make an M&E system work. In the third section, **Institutional framework**, the whole issue of how to strengthen the government M&E system is discussed. It is suggested that the starting point is to carry out a diagnostic study of what exists already and in the process to identify the level of interest and commitment to M&E of the different stakeholders. The development of a national M&E system requires substantial effort and a high level of central coordination as well as the active cooperation of sector ministries and agencies. Consensus is not easy but the likelihood of success is significantly increased if one is able to identify at the highest level, an M&E champion, who can push to make public services more results-oriented and thereby justifying the reason for having an M&E system at all. A second key ingredient is the coordination of the many different players involved. Country experiences would seem to indicate that the establishment of a central M&E coordinating unit is essential. The coordinating unit should be linked to national body that is responsible for overseeing the implementation of national strategic plans such as poverty reduction strategies. Its functions would probably include the preparation of an annual national M&E report, the establishment of national reporting standards, and the commissioning of evaluation studies. It would also have a quality assurance role with respect to the activities of the different sector M&E units. The preparation of the diagnostic study should lead to the setting up of an M&E capacity building action plan and would almost certainly include a significant level of support the strengthening of the national statistical system.

The sourcebook finishes by looking to the future and speculating on how M&E is likely evolve. It notes how, rather than being a marginal academic exercise, M&E he has started to emerge as a key agent of development in its own right, and an essential component of poverty reduction strategies. In-country demand, which has hitherto been quite limited, is growing and there is the beginning of a recognition that M&E information is not just a tool for policymakers and planners but a tool for members of the public and for civil society to hold public sector managers accountable. In this way, M&E systems are starting to become tools for promoting good governance and accountability.

Secondly it looks at how, in more and more countries, the strengthening of subnational administrations is becoming increasingly important development. The rationale of this is that decentralized decision making can bring governments closer to the people. However in many countries the technical capacity of government departments at subnational level is extremely weak, thus a major programme of capacity building is required in all areas including M&E.

The final issue concerns the monitoring and evaluation of global targets such as the MDGs at international level. The challenge for the international agencies responsible for maintaining cross-country databases is how to make the data submitted from disparate countries comparable with each other. The quality of many international statistical series is woefully low due to the fact that the individual country submissions are often of low quality and filled with gaps. They are in many ways basically incomparable because different methodologies and standards have been applied in the different countries. One of the long-term goals at international level must be to encourage a process of alignment to improve the compatibility of the different submissions from each country. This may be a long time coming but a first step in this direction has to be to address the issue of introducing standards for the metadata (including all the extra information about how the data were collected, how reliable they are, sample size etc.)which accompanies the data. Good metadata will very significantly simplify the task of the compilers at the international level and produce better and more comparable statistical indicators for monitoring at the global level. The sourcebook suggests that the relationship between national and international systems is not a hierarchical one, but is complex and symbiotic with the international institutions needing the outputs from the national institutions as much as the national institutions need inputs from the international institution. Ultimately, the global M&E network is only as strong as its weakest link. So even if for no other reason than to maintain the standard of international reporting systems, the donors have a vested interest in seeing that the capacity of national institutions is strengthened.