

PN-ABD-224
62401

Workshop on Improving Agricultural
Research Organization and Management:
Implications for the Future

ISNAR, September 8-12, 1986

EVALUATION OF AGRICULTURAL RESEARCH

by

MATTHEW DAGG

Senior Research Officer

ISNAR

EVALUATION OF AGRICULTURAL RESEARCH

Introduction

Evaluation is a very broad word, and it is used to mean a wide variety of operations. Even evaluation of agricultural research has a wide range of interpretations; so wide, that a discussion on one aspect of it can sound quite foreign to others concentrating on another aspect. At a recent meeting in Singapore on the role of evaluation in agricultural research, for some countries evaluation meant primarily the ex-ante examination and approval of research proposals (sometimes required by law before release of funds); for others (again required by law) it meant the review of technical, physical, and financial progress of research institutions and staff performance; for yet others evaluation of agricultural research meant little other than the assessment of the research contribution in the macro impact of agricultural development on the economy; and for some, evaluation meant what donors required of them in relation to projects involving agricultural research. The methodologies and approaches required for these different operations are clearly variable, and the uses of the information gathered also differ. However all are quite valid uses of the term evaluation.

What then does ISNAR mean by evaluation of agricultural research? There are two main aspects. Firstly, ISNAR is heavily concerned with strengthening national agricultural research management, and therefore must be concerned with evaluating agricultural research management systems at various levels in countries. It has had to develop its own methodology for carrying out reviews and evaluations of agricultural research systems management, so as to assist in developing

recommendations and programs for strengthening agricultural research. A "Guidelines for ISNAR Reviews and Evaluations" has already been published as a basis for such operations.

Secondly, ISNAR is concerned with training and offering advisory services on specific aspects of research management. Monitoring and evaluation are essential components in the management of any system: they provide mechanisms for the feedback of information to management on the progress of operations so that any necessary adjustment or corrections can be made to maintain or improve progress. Monitoring and evaluation of progress in the system are important in the integrated management of agricultural research, and ISNAR is therefore aiming to develop a comprehensive understanding of monitoring and evaluation in research in developing countries. The management system for national agricultural research is quite complex: different management mechanisms, involving different people, have to be in place to help determine major priorities; to decide on broad programs and to decide on annual programs for research; to develop a well trained capacity for research; to provide and control funds and facilities to carry out the research; to communicate the interpreted conclusions from research to a range of clients; and for all these, management must have feed back information on the progress and accomplishments of the research program so that plans and programs, and also management mechanisms, can be modified accordingly. Monitoring basically provides information on whether the program is proceeding as planned. Evaluation in general examines why the program is departing from the plan (and how it can be corrected, if necessary); and also whether the plan and organization are the best ways to achieve the overall objectives and goals. This applies at all stages in the system, from the researcher evaluating proposals and progress in an experiment,

through an institution checking its effectiveness, right up to the cabinet reviewing returns to its investment in research (see later, Fig. 2).

In its concern for in-depth understanding of the major components of agricultural research management, ISNAR must encompass the principles and appropriate methodologies involved at all stages of evaluation of NARS. A draft ISNAR working paper on concepts, organization, and methods for "Evaluative Activities for Management and Planning of Agricultural Research" has already been prepared by Dr. J. Murphy, which has proved very valuable in training seminars, even though it does not cover yet all of the stages of evaluation. It has also helped in guiding assessment of performances of component institutes in a research system.

There is already a vast literature on monitoring and evaluation and it might be wondered what more is needed: the principles and essential methodologies have been well expounded. However, much of the discussion has been on the monitoring and evaluation of the production and sales sides of business and development rather than on research. Similarly in agriculture, most of the literature is on evaluation of agricultural development projects involving extension, delivery of inputs, infrastructure, irrigation, credit, marketing, pricing, etc.; and also research as part of development process, but little special attention has been given to evaluation of research. There are however some important differences between the nature of agricultural research institutions and production and sales support institutions, that give rise to important differences and emphases in the methodologies used for evaluation, and these need to be taken carefully into account in considering what methodologies to use. For instance,

* Research creates only a potential for development. Its realization depends on many subsequent factors outside research. The immediate beneficiaries for agricultural research are development institutions and planners, not farmers or consumers. The latter, nevertheless, are the ultimate beneficiaries but often well down the line, and substantial increases in production and other economic effects from research advances commonly have a time lag of 10 - 20 years. For guidance of research management of institutions, evaluations with relatively short term indicators are necessary.

* Efficiency of research frequently depends more critically on the determination of the research program than on its implementation, while development agencies lay great stress on physical and financial efficiency of project implementation.

* Development projects can often be planned in detail and positive progress can be related to physical achievement on a planned time frame. Research is intrinsically uncertain in its timing of progress, and the short term research program must be continually flexible, depending on the nature of current experimental results. Indeed, individual researchers are constantly monitoring their own progress and making important small scale management decisions in the normal course of their operations. Moreover, a lot of good research work must yield essentially negative conclusions (e.g. nothing is better than current recommendation), which is difficult to accommodate in a framework geared to positive progress.

The differences are great enough to merit special attention to methodologies for evaluating agricultural research. This is being recognized; for instance, no attempt is now made to calculate economic rate of return ex-ante from World Bank assisted agricultural research projects.

Some of these differences will be considered in the next section which reviews briefly some principles and methodologies for evaluations of agricultural research, and the following section will consider aspects of particular concern to ISNAR.

Some general principles of evaluation and methodologies

As indicated earlier monitoring and evaluation are integral components of an on-going management system as a means to draw lessons from past experience and to incorporate them into the planning and implementation of future activities. They have not always been seen as such in developing countries. The introduction of separate monitoring and evaluation units in association with external donor assisted agricultural development projects left an unfortunate impression of surveillance and judgement from outside as far as the country project staff were concerned. However, from the donor viewpoint, the operation was an eminently sensible feedback of information on project performance to the management of the donor aid agency to guide changes of plan and design of future projects. To the larger organization, the information collected on the performance of a unit within it looks like a monitoring exercise on a long time scale; viewed from within the unit, the same exercise looks like a one-off evaluation of competency and efficiency of staff performance. Size, viewpoint and time scales are important parameters in monitoring and evaluation.

Similarly, what is essentially a continuous and continuing management operation on the large scale, can be broken down to discrete time bound projects on smaller scales. Much of the literature and methodologies are indeed concerned with project evaluation, and though time bound projects are always integral parts of a larger continuous system, the project gives a specific and convenient base for assessing progress and performance.

Definitions

A United Nations task force on rural development has provided appropriate definitions of monitoring and evaluation: "Monitoring is the continuous or periodic review..... by management at every level of the hierarchy of the implementation of an activity to ensure the input deliveries, work schedules, targetted outputs and other required actions are proceeding according to plan".

"Evaluation is a process for determining systematically and objectively the relevance, efficiency, effectiveness, and impact of activities in the light of their objectives. It is an organizational process for improving activities still in progress and for aiding management in future planning, programming, and decision making".

Monitoring essentially involves comparison of observed performance with a plan or program. Evaluation involves a wider range of comparisons illustrated in figure 1. This is most readily appreciated relative to evaluation of a program or project, but can be applied more widely to an on-going system.

In figure 1 the first comparison, before/after, is historical assessment. The with/without comparison helps understand whether the program has made any difference. Differences are always obscured to some extent by "noise" in both situations due to changes occurring due to factors not controlled by the program. The third comparison, achieved/expected, compares what has been done with what had been planned or expected. It can give an evaluation of whether the program's implementation has been satisfactory; or it can provide indications that some of the assumptions in the planning were not fulfilled.

Figure 1: Possible Comparisons

	T ₁	T ₂	
outside program	Y ₁	Y ₂	
			with/without
			before after
program achieved	X ₁	X ₂	
			achieved/expected
program planned	X ₁	X'	

- compare situation X at time T₂ with what it was at time T₁;
- compare situation X at time T₂ with situation Y, where no program existed;
- compare situation X at time T₂ with an expected X' which was defined when the program was designed in the first place.

This, of course, is a very simplistic representation. Real evaluations bring in great complications, especially for instance, in evaluating the performance of a research institution. But the basic comparisons remain valid, even if non-existent plans and predictions have to be reconstructed, or performance norms invoked to be a basis for comparison. However, this simple picture does emphasize that for new programs and projects, evaluation must begin when the program is being designed, by setting clear objectives for the program and by selecting specific, verifiable indicators of achievement, and by specifying how the achievements will be measured. This provides the basis for monitoring procedures, and systematic monitoring builds up a data base that provides essential information for subsequent evaluation.

Types of evaluation (objectives)

The objectives for introducing a program or project are many and varied, but those ultimately responsible for its management and success will expect information from evaluations on an on-going or completed program about the efficiency, the effectiveness, and the impact of the program. These are the basic types of evaluation, and they require quite different methodologies in data collection, and different indicators. (Relevance is another candidate for classification as a type, but can perhaps be subsumed in efficiency as far as agricultural research is concerned.)

Categories of evaluation (timing)

The basic comparisons are not necessarily only made after a program is completed in real time. Evaluations of the predicted results of

programs are the basis of the pre-appraisal of projects and programs, (potential impact), and indeed the basis for setting out the planned performance that will be monitored. Furthermore, interim evaluations during the course of a program provide valuable information performance for modifying the plan. There are therefore three basic categories of evaluation: ex-ante, monitoring, and ex-post. Reviews and evaluations of the performance of on-going institutions can be a mixture of all three categories.

Response time

As mentioned earlier, evaluation activities are needed at many management levels of an organization or system, to provide feedback information to guide modification in planned operations. A basic principle is that the feedback information shall be delivered rapidly enough to prevent massive departures from performance objectives. It follows that the response time of a performance indicator must be short compared with the characteristic time of operation of the system component involved.

The frequency at which an evaluation is carried out depends on how tight a control on the operational process is required. Usually the time interval will be longer than the response time of an indicator, but this is not necessarily so. Conclusions from evaluations can refer to effects of actions taken before the previous evaluation but which had no chance of expressing themselves before the previous evaluation.

Methodologies

Different types and purposes of evaluation require different methodologies. The numbers of examples start getting rather large but some generalization is possible. The best methodologies for evaluation are solidly founded in good monitoring, in a time series of well chosen observations. The management of all projects and on-going systems require feedback from routine monitoring of physical and financial progress, data from which will be an important help in periodic evaluations of efficiency. General methods for this exercise are very well known, and the project budgeting system, so valuable in research program formulation and management, can be modified to incorporate this important monitoring activity.

At the macro agricultural development level, impact evaluation on ultimate beneficiaries will require collection of information on macro increases in production, and/or incomes, exports, nutrition, etc., (even political shifts). The methodology involves socio-economic surveys and sampling techniques, which are well understood but difficult, costly and time consuming to carry out satisfactorily. These are quite different techniques from those employed in socio-economic and technical analyses carried out in ex-ante evaluations of potential impact.

Methodologies for other types at different management levels can vary widely. Two general techniques are particularly useful. Beneficiary contact monitoring involves testing project products on small numbers of target beneficiaries at an early stage to get some idea on potential impact (e.g. on-farm research). Diagnostic studies are similar; it

involves mounting small supplementary studies to test hypotheses about why performance is deviating from plan. It is excellent for rapidly discounting hypotheses.

In addition there are a range of more subjective review procedures to assess technical contributions to potential development, and effectiveness of organizational arrangements.

All of these methods need appropriate and sensitive indicators (as few as possible to keep measurement and analytical costs down) that reflect the progress desired and have a response time short enough to allow timely modifications of plans and programs.

Application to the national agricultural research system

Most evaluations of agricultural research have been carried out on donor-assisted projects (with notable exceptions in a few countries, e.g. India), and they have usually been carried out within a donor agency management system to provide information to improve its overall operation, with requirements and values that are not always those of the national research management. ISNAR is essentially concerned with improvement of the management component of monitoring and evaluation within the national system.

There are management groups at several levels within the national research system that would benefit from an improved information feedback system on its performance. Some are listed in figure 2. Each management level will have different expectations from evaluative exercises carried out, and different uses for the findings. They will need different types

Fig. 2: Expectations from evaluations of agricultural research at different levels of management in the system

Level of management	Expectation from information from evaluation	Type	Category	Response time	Methodology	Criteria (Type of indicator)	Frequency	Use
1. Cabinet	Use of investment in agric. research	Impact	Ex-post	10 yrs +	Socio-economic survey	Macro-economic, political	5-10 yrs?	Guide investment level
2. National agricultural ministries	Use of investment in agric. research	Impact	Ex-post	10 yrs +	Socio-economic survey	Macro-economic, political	2-5 yrs	Guide balance of investment in research/development institutions
		Effectiveness	Ex-post	5 yrs	Technical Review	Technical contribution/Development relevance		
3. National agricultural research (Council)	Effectiveness of research institutions;	Effectiveness	Ex-post	5 yrs	Technical review, Diagnostic studies	Technical contribution/Development relevance, Costs, Management mechanics	1-5 yrs	Guide allocations to research institutions;
	Output to development institutions	Efficiency	Monitoring	2-3 yrs	Physical/financial/organizational; Beneficiary monitoring			Improve efficiency of management of research institutions;
		Potential impact	Ex-ante	?	Technical and socio-economic analysis			Help review schemes of service

1/10

Fig. 2 continued (page 2)

Level of management	Expectation from information from evaluation	Type	Category	Response time	Methodology	Criteria (Type of indicator)	Frequency	Use
4. Research institutions	Effectiveness of management components in operation of inst.;	Effectiveness	Ex-ante Monitoring Ex-post	2-5 yrs	As in 3 above plus Annual research plans and reports;	Technical contribution/ Development relevance, Costs,	1-5 yrs	Guide allocations to divisions;
		Efficiency		0.3-3 yrs				
		Impact, Potential impact	Ex-post Ex-ante	5 yrs ?	Personnel review	Management mechanics, Requirements of people, Linkages		Long term research program; Assessment of personnel; Argument for resources; Improvement of management mechanisms
5. Research station	Research performance in relation to budget, facilities and program; Relevance of research experiments; Identification of constraints	Efficiency	Monitoring	0.1-1 yr	Physical/ financial/ organizational;	Efficiency of experiments, Costs,	Annual	Guide allocations to stations;
					Beneficiary monitoring;	Use of personnel and facilities		Short term (annual) program;
					Annual review and reports; Personnel review			Personnel management

111

Fig. 2 continued (page 3)

Level of management	Expectation from information from evaluation	Type	Category	Response time	Methodology	Criteria (Type of indicator)	Frequency	Use
6. Researcher	Quality and output of research;	Effectiveness	Monitoring	1-3 yrs	Personnel performance review;	Quality and quantity of experiments and studies;	1-5 yrs	Guide to short term program;
	Management of resources;	Efficiency	(and ex-post?)		Annual review and reports;	Perception of task and role;		Staff development
	Assessment for promotion				Program proposal review;	Responsible performance;		
						Linkages and communication		
7. Project	Progress in relation to project plan and objectives	Effectiveness	Ex-ante	(Ex-ante?) 2-3 yrs	All (except personnel review, and macro socio economic survey (?)) plus benefit/cost analysis	(Depends on project)	Mid term, Terminal, Ex-post	Guide to modify project;
		Efficiency	Monitoring (terminal)	0.2-1yr		Costs;		Assessment of value, impact;
		Impact	Ex-post	10 yrs +		Use of funds		Guide for future projects

and categories of evaluations. A first attempt is made in figure 2 to fill in some of the matrix to indicate the complexity of the situation, but also to indicate some of the limits that can be placed on the need for evaluative information at different levels. The levels of management are not, of course, isolated and appropriate findings from evaluations carried out at any level (or combined operations) would be passed up and down the system.

Discussion from participants is invited at each and every point. Least detail is ventured in the column for indicators. Clearly there is a lot of scope for deep and creative thought about efficient indicators that genuinely reflect desirable progress, and not just the physical and financial performance (important though that is). ISNAR hopes to gather and disseminate examples of good indicators for monitoring and evaluation for different management situations.

Salient features of figure 2 are the steady shift from macro to more micro concerns, with attendant need for shorter response times in indicators, and a shift from evaluations of impact to evaluations of effectiveness and efficiency. (The response time for effectiveness at Research Council level reflects the time a research conclusion may take to be fully accepted by development institutions - the immediate beneficiary and the means of reaching large numbers of producers.) Ex-ante evaluations of potential impact is not given a response time: conclusions from technical and socio-economic analysis (even though these may take some time to carry out) are presumed to go straight into action in the form of new projects and studies.

Some particular aspects of evaluation of agricultural research

Some important differences between the role and nature of research in development as compared to other development institutions were mentioned earlier. They carry through to significant differences in evaluation methodologies and must be taken into account.

Macro socio-economic survey: To repeat perhaps the most important point: research generates potential and new opportunity for development; but development institutions must be convinced and mobilized to realize that potential. Macro impact must be demonstrated at the ultimate beneficiary level (producers and consumers) to impress government in development terms of increased production, income, nutrition, etc., and the research contribution towards this (of 10-20 years earlier), extracted by analytical process on data collected through samples and surveys. This is a vital step for convincing Government to allocate more funds to research, but not necessarily a good guide for changing current research management or current research priorities and plans. An indicator with shorter response time is needed. For instance, an appropriate group of indicators to convince a Research Council of effective impact could be evidence of development institutions beginning to take up new technology proposed from research.

However, the farmer remains the most important ultimate beneficiary of research, and his response to new technology is also critical. Assessment cannot wait until the technology is delivered years later via extension and other agencies. On-farm research is an excellent opportunity for implementing Beneficiary Contact Monitoring - to monitor that the research program is making progress and to evaluate whether or

not it is relevant, in the right direction to benefit the farmer. Promising technological improvements can be tried out with a few farmers even before research has finished polishing the technology to the stage of recommendation to development institutions. (Of course, extension should be part of this on-farm research exercise). The chance of verifying the result through production increase on a district scale would be many years away, and would be a difficult and costly operation, even if permanent sampling teams were in place.

Except for very occasional macro-impact evaluations to impress Government on its role in total production, agricultural research probably has no need of evaluations involving large scale sampling exercises, which is where most of the difficulties in evaluation methodology lie. This would make prospects of evaluation of agricultural research appreciably easier.

Organization and management: Efficiency of performance in national research depends heavily on the organization of the research institutions in both determining a very relevant research program and carrying it out. Evaluation of the performance of the organization is therefore important, especially in projects concerned with institution building, and suitable indicators must be sought. It is, however, difficult to find appropriately sensitive indicators with response times as short as a few years. In order to review and evaluate research institutions with a view to improving its organization on a medium time scale, it is therefore often necessary to use as a basis for comparison, norms for probable good performance, based on comparative experience of research organizations elsewhere and a clear appreciation of the nature and management of agricultural research.

The need for some "model" of how a national agricultural research system, or components, work in supporting development, is also necessary for predicting the course of a plan or project - the project intervention model that must underlie project proposals. This is again especially important on projects concerning changes in organization and institution building. ISNAR is continually working towards a deeper understanding of how components of the management system of agricultural research must integrate consistently into a coherent framework based on national policy. (For instance, it is likely to be counter-productive to have a reward system for researchers based on international publications when national policy is to boost small farmer production and for the research institution to ensure that they have the best adapted technology.)

Program budgeting system: As in any organization, it is essential to have feedback information on the performance of physical and financial progress, and this must have a substantial place in monitoring and evaluation of research. However, in agricultural research it is particularly important that the physical and financial progress should be closely linked to a carefully determined program. Moreover, monitoring must be done in comparison with a detailed planned program. The process of program determination discussed earlier by Dr. Devred is valuable for fixing the annual program of work; and the program budgeting method is excellent for defining the path of physical and financial progress to facilitate the program. A program budgeting exercise of this kind is essential for any efficient monitoring and evaluation of physical, financial, and program progress.

Summary and conclusions

Monitoring and evaluation are essential components in efficient management of national agricultural research systems to provide feed back information on how to modify plans, programs, and organizations.

Evaluation of agricultural research is quite different from evaluation of development projects in several characteristics, and needs a significantly different approach and emphasis in methodologies.

Within the hierarchy of management levels in the national agricultural research system, there are a wide range of users and uses of evaluative activities, requiring different types of evaluation, methodologies, and time scales for response.

Future work of ISNAR will focus on identification of appropriate and sensitive indicators of research performance that will give efficient feed back of information to different levels of management to permit modification of operations to improve the efficiency of the research system.