WORKING PAPER No. 14

PN-ABA-230

MONITORING AND EVALUATION IN THE MANAGEMENT OF AGRICULTURAL RESEARCH

Diana McLEAN

isnar

International Service for National Agricultural Research

The International Service for National Agricultural Research (ISNAR) began operating at its headquarters in The Hague, Netherlands, on September 1, 1980. It was established by the Consultative Group on International Agricultural Research (CGIAR), on the basis of recommendations from an international task force, for the purpose of assisting governments of developing countries to strengthen their agricultural research. It is a non-profit autonomous agency, international in character, and non-political in management, staffing and operations.

Of the 13 centers in the CGIAR network, ISNAR is the only one that focuses primarily on national agricultural research issues. It provides advice to governments, upon request, on research policy, organization and management issues, thus complementing the activities of other assistance agencies.

ISNAR has active advisory service, research and training programs.

ISNAR is supported by a number of the members of CGIAR, an informal group of approximately 43 donors, including countries, development banks, international organizations and foundations.

WORKING PAPER No. 14

MONITORING AND EVALUATION IN THE MANAGEMENT OF AGRICULTURAL RESEARCH

Diana McLEAN

May 1988

isnar

International Service for National Agricultural Research

ISNAR WORKING PAPERS

The ISNAR working papers series is a flexible instrument for sharing analysis and information about relevant organization and management problems of the agricultural research systems in developing countries.

In the course of its activities - direct assistance to national agricultural research systems, training, and research - ISNAR generates a broad range of information and materials which eventually become the formal products of its publication program. The working papers series enhances this program in several important ways:

1. These papers are intended to be a rapid means of presenting the results of work and experiences that are still in progress, but are already producing results that could be of use to others.

2. They are intended to be an effective vehicle for widening the discussion of continuing work, thereby increasing the quality of the final products. Critical comment is welcomed.

3. The series provides an outlet for diffusing materials and information which, because of their limited coverage, do not meet the requirements of "general audience" publication.

The series is intended mainly for diffusion of materials produced by ISNAR staff, but it is also available for the publication of documents produced by other institutions, should they wish to take advantage of the opportunity.

MONITORING AND EVALUATION IN THE MANAGEMENT OF AGRICULTURAL RESEARCH

TABLE OF CONTENTS

	Page
Preface	
Introduction	2
Types of Research Program Evaluation	5
Research evaluation: performance, quality, and relevance	5
Ex ante evaluation	7
Evaluation of ongoing research	8
Ex post evaluation	9
Impact evaluation	10
Evaluation Criteria and Methods	11
Setting Up a Monitoring and Evaluation System	14
What are your M/E requirements?	14
Is a separate M/E unit needed?	15
How much does M/E cost?	16
Guidelines for evaluation	16
ISNAR's Approach to Developing M/E Materials	18
Annex 1. Monitoring and Evaluating Ongoing Research	19
Definitions of terms and concepts	19
Relationship between ex ante and ongoing research evaluation	19
Monitoring ongoing research	20
Evaluating ongoing research	24
Role of peer and expert review	25
Role of strong program leadership	25
Annual project and program evaluation	25
Comprehensive program evaluation	26
Summary	27
References	29

PREFACE

This paper was prepared to introduce agricultural research managers to the basic concepts and approaches of monitoring and evaluation (M/E). It provides the framework for the development of a series of materials at ISNAR on the comprehensive topic of monitoring and evaluation. These subsequent papers will attempt to provide research managers with techniques of M/E for planning purposes (ex ante), ongoing project and program evaluation, final project evaluation, and impact evaluation.

This paper introduces the general topic of monitoring and evaluation, including a brief definition of terms, and the functional roles of different types of evaluation in research systems. The importance of expert or peer review is discussed as the most fundamental method of research evaluation at all stages of planning and implementation. The paper underlines the importance of integrating monitoring and evaluation into routine management practices, so that they are viewed by both those conducting evaluations and those being evaluated as tools for improving research. In addition to the main text there is an annex which more thoroughly discusses the evaluation of ongoing research, largely through annual reviews and comprehensive program reviews.

This paper discusses the distinct yet interrelated aspects of evaluating research performance, quality, and relevance. It does not attempt to cover the monitoring and evaluation procedures associated with personnel appraisal and financial and administrative management. These topics are considered in other ISNAR papers on human resource management and in general management literature.

It is necessary here to define the terms "program" and "project", since they have different meanings in different research organizations. Programs are coordinated research activities whose combined scientific output addresses national research objectives. Programs are long-term and somewhat continuous, and are composed, in some cases, of sub-programs, and of projects. Projects address specific research problems, and have explicitly defined timeframes, resources, and targets. Each project in turn comprises a number of specific operations or experiments.

Finally, I would like to acknowledge the contribution of ISNAR staff and others to this paper. Josette Murphy, formerly of ISNAR, made substantial progress on the topic, reflected in the draft publication Monitoring and Evaluation in Agricultural Research: Concepts, Organization, and Methods, and ISNAR Working Paper #2, Using Evaluations for Planning and Management: An Introduction. Matthew Dagg and Peter Goldsworthy have worked on integrating evaluation with planning, and have begun attacking the problem of identifying appropriate indicators at different levels of technical programs. Each has served as the ISNAR working group chairperson for monitoring and evaluation. Other ISNAR staff have offered valuable support, particularly Ghazi Hariri, Genevieve Michel, Dennis Wood, Paul Marcotte, and Rudolf Contant. In addition, I would like to offer deep gratitude to S.A. Adetunji, A.M. Macha, D.R.B. Manda, S.N. Maturi, J. Nya Ngatchou, B. Patel, E. Whingwiri, and Yohannes Kebede, who served as expert consultants on this topic.

INTRODUCTION

Information is the most valuable input for decision making. In national agricultural research systems information is needed at all levels of planning and implementation: by policymakers, national planners, and agricultural research leaders. Processes of monitoring and evaluation (M/E) are the primary means of collecting and analyzing information, and are thus intrinsic to good research management. Over the past decade, research managers have become increasingly aware of the importance of installing M/E procedures into their organizations, but the successfulness of these efforts has been mixed.

There are several reasons to integrate M/E into national agricultural research systems. One major reason is that since the economic recession of the 1970s research funds have been limited, and governments have required more justification for and controls over funds. In addition to the recession-induced constraints, the expanding fields of science and the growth of research institutes have created a greater demand for the existing resources. In order to remain productive in the increasingly more specialized areas of science, intellectual leadership is needed in each field, and funding must be available for the most promising and important research areas. It therefore becomes necessary to evaluate exploitable areas of science, the research most likely to provide new technologies to fuel national development, and to identify areas of investment. In addition, evaluation can improve ongoing research, through improved technical decision making and management.

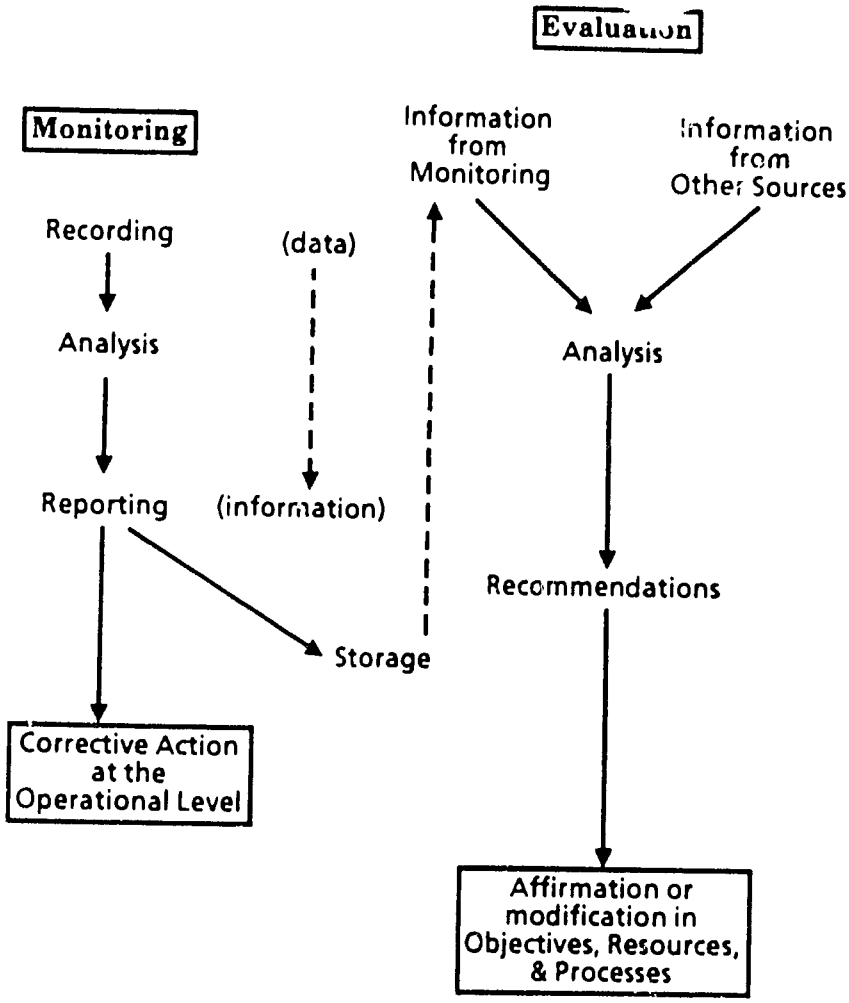
Monitoring is essential for evaluation (Figure 1), and evaluation is used to assess:

- the potential impact of research in priority-setting and planning exercises;
- the performance and quality of research in progress;
- the successful completion and relevance of research projects;
- the ultimate impact of research results on the achievement of development objectives.

Monitoring includes the periodic recording, analysis, reporting, and storage of data on key research indicators. Monitoring primarily provides information on project performance, on whether an activity is proceeding according to plan. However, if ex post project evaluations or impact evaluations are to be held, it will also provide information on socioeconomic indicators for these purposes. For ongoing research, managers primarily monitor resources, such as the use of funds and personnel, and processes, such as the occurrence of annual review meetings or periodic seminars. This permits management to compare the progress of work against planned objectives, detect deviations, identify bottlenecks, and take corrective action while research is in progress.

The best key indicators of project performance are objective, quantifiable, and unambiguous. They can be verified if necessary. A good monitoring system is not more time consuming than the benefits justify, collects no superfluous data, is timely in data analysis, interpretation, and feedback, and is useful to researchers.

Figure 1: Relationship of Monitoring to Evaluation



Evaluation is based on both qualitative and quantitative information, gathered through monitoring and from other sources. Whereas monitoring tracks research performance, whether progress is according to plan, evaluation analyzes issues of quality and relevance, and may even analyze the appropriateness of the plan itself. Evaluations result in a set of recommendations, which may address issues of planning, such as a shift in program objectives or content, or implementation, such as the need for more laboratory capacity.

Monitoring and evaluation are not new concepts. Yet research institutes have had little success in integrating effective M/E into their organizations. Why? Many professionals find it difficult to apply stricter management and control measures on what is essentially a creative process. They have come to accept that research has characteristics which distinquish it from most non-research activities, where M/E is more easily applied. Research is also intrinsically uncertain in its timing and its products, since progress depends on previous experimental results and breakthroughs made elsewhere.

The failure to conduct effective M/E can also be attributed to the attitudes of the researchers themselves. Monitoring and evaluation is regarded by many researchers as a burden inflicted upon them by bureaucrats, a process that leads from "analysis to paralysis". This attitude has been nurtured by the limited use of evaluation for inspection, auditing, or control purposes. In order for M/E to be used in a more positive manner, management and staff must have a common understanding of the importance of the processes involved, and of the contribution M/E can make to achieving the objectives of the research system. This requires that management fully support the integration of M/E into day-to-day operations.

Part of the problem also lies in the imprecise nature of the terms "monitoring" and "evaluation". The depth of analysis associated with each, their interdependence, and the use of the term evaluation to describe entirely different management processes have confused many managers. Any assessment, appraisal, analysis, or review is in the broader sense evaluative. However, in defining evaluation as a management tool, we have assumed that certain basic criteria exist:

1. There must be clearly defined targets and key indicators, and they must be determined in advance. This ideally occurs during ex ante evaluation, where baseline data and assumptions about project progress are explicitly laid out, and from which systematic monitoring procedures are set up. The data monitored often take the form of a time series of well-chosen observations.

2. The principal purpose for carrying out an evaluation needs to be clearly stated. It should be kept in mind that research does not function in isolation, and therefore evaluators must consider the political, institutional, social, and economic context in which research is conducted. Evaluation occurs at different management levels for different purposes, and can be concerned with such diverse factors as the use of institute resources, the scientific quality of research, the appropriateness of the technology, or national returns on investment in different sectors.

3. The scope of the evaluation must be clearly understood. When evaluating the impact of research, a differentiation must be made between research results and the contribution they make toward greater development objectives. Since this contribution is influenced by national infrastructure, economic parameters, extension possibilities, etc., successes or failures in agricultural development cannot be ascribed solely to technology generated from research. Research creates only the <u>potential</u> for development; whether or not this is realized depends on many other factors.

TYPES OF RESEARCH PROGRAM EVALUATION

Information from evaluations is used in different stages of research management (Table 1), by all levels of the management hierarchy. It is used in the management of technic l programs, personnel, and financial resources. All research managers have participated in some monitoring and evaluation exercises, such as personnel appraisal, program planning, or end-of-project review of donor-funded activities. But most research institutes do not use M/E to maximum advantage. This section discusses the types of evaluations that are useful to managing the research program itself.

Part of the confusion surrounding the discipline of evaluation is an inconsistency in the terminology used. This becomes even less clear in the area of research evaluation, where little previous work has been done. In order to discuss the concepts, issues, and methods of the different types of evaluation, it is necessary to define a working vocabulary. Since there is no consensus on the terminology, ISNAR has accepted the following broad definitions.

Research evaluation: performance, quality, and relevance

The evaluation of technical research programs involves three important aspects. The performance of the research program compares achieved with expected outputs. It is primarily concerned with the use of resources, and the timeliness of the activity, and is determined mostly through monitoring and ongoing evaluation. Results from performance M/E are used to improve management procedures and increase productivity.

Assessing the success or failure of research, however, goes far beyond determining whether resources were used according to plan. Quality in research execution - the adherence to accepted standards of scientific work and precision - is essential. Evaluating quality in research poses serious problems for research organizations that lack experienced scientific expertise to cover all the disciplines and areas being researched. The quality of research is determined almost exclusively through some form of peer or expert review. Due to its importance and the problems encountered with peer review in many organizations, the topic will be covered in some detail in this paper.

Finally, research must be relevant. In most developing countries, research is "mission-oriented"; that is, it attempts to solve constraints or provide opportunities to national development. Evaluating the relevance of research involves relating each level of research objectives (from project to program to institute) to the next higher objectives, which ultimately reflect development objectives. Relevance, too, is primarily assessed through peer or expert review; in this case expanded from a more specific disciplinary focus to include experts in social, economic, and even political sciences. Determining the relevance of research often involves feedback from on-farm research and data collection through formal and informal surveys.

To discuss evaluation, it is necessary to distinguish between types of evaluation, since practitioners use terms inconsistently. Evaluations are most often categorized according to when they occur and their

Table 1:

Roles of Monitoring and Evaluation at Different Levels of Management in Agricultural Research Systems

Level of management	Type	Methods	Frequency	Use
1. Cabinet	Impact Ex ante	Socioeconomic survey Technical & socio-economic analysis	10-15 yrs 3-5 yrs	Guide investment level to broad sectors
2. Agriculture ministries	Impact Ex ante Ex post	Socioeconomic survey Technical and socio-economic analysis, and review	10-15 yrs 3-5 yrs	Guide balance of investment in research/development institutions
3. National agricultural research (council)	Ex ante Ex post	Technical and socioeconomic analysis and review	3-5 yrs	Determine potential impact of research initiatives to guide allocations to research institutions Verify original assumptions/lessons learned
4. Research institutions	Ex ante	Technical and socioeconomic analysis	3-5 yrs	Determine potential impact of research initiatives; justify/allocate resources to divisions/programs
	Monitoring	Resource/process evaluation	periodic	Improve efficiency of management of research institution
	Ongoing	Peer/expert review	annual	Improve research implementation and planning
	Ex post	Technical and socio-economic review	3-5 yrs	Lessons learned: comprehensiveness & complementarity of programs
5. Research stations	Monitoring	Resource/process evaluation	periodic	Improve station management
6. Programs	Ex ante	Technical & socio-economic analysis	3-5 yrs	Determine potential impact of different research approaches, projects
	Monitoring	Resource/process evaluation	periodic	Improve program management
	Ongoing	Peer/expert review	3-5 yrs	Guide short-term program planning
7. Projects	Monitoring	Resource/process evaluation	periodic	Improve project management
	Ongoing	Peer/expert review	annual	Guide to modify projects in progress
	Ex-post	Technical/socio-economic review	3-5 yrs	Guide for future projects
8. Researcher	Monitoring	Resource/process evaluation	periodic	Improve activity management
	Ongoing	Peer/expert review	annual	Guide research execution and planning

purposes. The evaluation of research occurs before (ex ante) to assess the the potential impact of research, during (ongoing) to evaluate the performance and quality of research projects in progress, immediately after (ex post) to determine the successful completion and relevance of research projects, and 10-15 years after research results have been achieved (impact), to assess its ultimate impact on development.

Ex ante evaluation

Ex ante evaluation is a research planning process, which includes a comprehensive analysis of the potential impact of alternative activities before implementation. The primary method used in ex ante evaluation is peer or expert review, and the fundamental techniques are the same as those described in the paper "Priority Setting in Agricultural Research": primarily checklists, scoring models, and benefit/cost analysis. Many people associate the term "ex ante evaluation" with economic analyses, often at the national planning level. However, ex ante evaluation can occur at all levels, in institutes, programs, and even research projects, and the term is often used interchangeably with what different people call priority setting, planning, and program formulation.

Different evaluation criteria are considered at these various levels, depending upon the analysis being conducted. For instance, at higher levels, where planning takes into consideration the allocation of resources to different commodities, economic data would be more important than at the program level, where scientific and technical criteria become dominant. Some of these various criteria are potential economic impact, compatibility with development goals and research objectives, urgency for research, resource availability, the appropriateness of research results, including ease and cost of adoption by farmers, political influences, regional concerns, and contribution to technical programs and world knowledge.

The techniques also vary. Techniques of cost/benefit analysis, and econometric techniques using a production function approach are sometimes used at national or institute levels of decision making. While these techniques are theoretically applicable, it is often difficult to obtain the necessary personnel and data to conduct the analyses, and it is hard to identify and analyze the externalities arising from research which add benefits not captured in cost/benefit models. A possible exception is very specific research, such as the development of hybrids, where benefits can be captured by a single indicator, say improved yield.

At the project and program levels the techniques most often used are checklists and scoring models, with less emphasis on economic analysis. For these more subjective techniques, expert advice becomes extremely important. Research organizations differ in how researcher oriented they are; that is, how much individual judgement influences scientific priorities. In more project-oriented systems, the reviewers have more control over the development of the research program. A positive benefit of the project system is that researchers must clearly formulate objectives and workplans. This makes management-by-objective possible, which can economize on scarce resources. The project approach also encourages interaction between experienced and less experienced researchers in the planning of research proposals.

Ex ante evaluations are not always as comprehensive as they should be. As currently practiced, they tend to be restricted in their focus to particular disciplines. For instance, an economist might conduct an ex ante evaluation to look at potential returns to investment. A rural sociologist, through diagnostic surveys, would enlarge the criteria to include social benefits from research. A biological scientist might look only at scientific merit or the potential to make a technical breakthrough. To make ex ante evaluations more effective, there should be participation from different disciplines, and more comprehensive criteria should be applied. Thorough ex ante evaluation defines the baseline against which progress will be measured, sets targets, and states the assumptions used in making the projections. In addition, in order to tie ex ante evaluation to ex post evaluation, the indicators to be monitored during the course of the activity should be defined.

Even at the highest levels of management, effective ex ante evaluation depends on information from the researchers in the system. Management must assure mechanisms for conveying feedback from researchers on what they consider necessary or good opportunities for research investment, the potential contribution of different lines of research to development, and current or projected bottlenecks preventing the full adoption and use of research results. Likewise, management must communicate priorities and national objectives to researchers so operational planning can take place at the program and project levels.

Evaluation of ongoing research

The evaluation of ongoing research, also called ongoing or interim research evaluation, is the periodic review of research in progress. It is used to analyze not only the use of resources but also the quality of research, and the continuing relevance of research programs and projects. Ongoing evaluations can be annual in the case of project reviews, on demand in the case of suspected problems in implementation (such as unforeseen budget cuts or the departure of key staff), or every three to five years in the case of a comprehensive program review. Ongoing evaluations also consider program processes, by looking at such things as the occurrence and quality of the annual project proposal and review process, or the administration of research stations.

The most prevalent form of ongoing research evaluation looks at research performance, comparing achieved with expected activities, usually on an annual basis in the case of research projects. It indicates how efficiently resources are used and identifies implementation problems. It is important to note that efficient research is not necessarily successful. The effectiveness or relevance of the research conducted, as defined by its original objectives, is of greater importance.

Accordingly, recommendations from this level of ongoing evaluation mainly address problems associated with the day-to-day management of research, although they may also indicate the need for changes in project objectives and targets. Monitoring is fundamental to this type of evaluation. It primarily tracks the provision and delivery of inputs and services, and requires the generation of information on the availability and deployment of staff, infrastructure, equipment, supplies, services, and funds for projects within programs. While not necessarily the

responsibility of project leaders themselves, a lack of awareness of the availability and quality of resources can have negative effects on the research in progress. If reporting systems for personnel and supplies are functioning reasonably well, and a proper work plan with defined resources and benchmarks has been prepared in the research proposal, it is relatively easy to install an adequate monitoring system for performance evaluation.

Apart from the day-to-day monitoring of resources, a distinctly different monitoring process is used in the execution of applied and adaptive research; that is, receiving feedback from clients - farmers, extension services, and development organizations - during the course of the research. This is most often done through on-farm research. This feedback in the early stages of experimentation can save substantial resources and effort if a reorientation is required.

Ongoing research evaluation is also concerned with the quality of research, and to some degree its continuing relevance. This assessment is most often conducted through expert or peer review under the technical leadership of heads of institutes or programs, and assures the soundness of technical decisions at all levels in programs and projects. The adherence of scientists to accepted research standards should be an important part of the annual project review. This is not always the case, where either time constraints, lack of experienced researchers to serve as peers, or an over-emphasis on performance exists. The issue of continuing relevance of program and project objectives is most often included in comprehensive program evaluations.

Ex post evaluation

An ex post or final evaluation assesses project performance, quality, and relevance immediately after project completion. It attempts to measure the effectiveness and efficiency of a completed activity, and includes an analysis of the original assumptions used in planning. Ex post evaluation might be confusing to those research managers who, erroneously, establish no time frames for projects. Using established norms, the results which a project is expected to achieve in a given time can be specified. Estimating the time frame is part of the ex ante evaluation exercise. Good ex post evaluation is linked to ex ante evaluation, and can best be conducted where a baseline was originally defined, targets were projected, and data were collected on important indicators. Unfortunately, the link between ex ante and ex post evaluation is poor in many organizations, and the two are often even conducted by completely separate groups. Ex ante evaluation is most often conducted through expert consensus, while ex post evaluations are increasingly becoming the preserve of professional evaluators. Care must be taken to avoid this separation.

Ex post evaluation is more in-depth than the annual project evaluation, and analyzes the project from beginning to end, determining whether project objectives were attained, causes for discrepancies, costs, and the quality and relevance of the research. Ex post evaluation indicates which projects might be continued, expanded, reduced in scope, or terminated. Lessons learned from ex post evaluation are incorporated into subsequent project planning exercises. Due to the costs of conducting ex post evaluations, they are usually done only for the more important, innovative, or controversial projects.

Ex post evaluations often consider such aspects as the cost effectiveness of research; its potential relevance to national development goals; the response of research to an urgent and important problem; the acceptance of results by farmers, extension services, and development agencies; and the contribution of research to world knowledge. The aspects considered and the depth of the analysis will vary with the importance of the activity and the interests and values of the organization, such as those having economic, social, or nutrition biases.

While indicators for research performance are easily identifiable, much more work is needed to identify evaluation indicators which deal with research quality and relevance. In most developing countries research is meant to serve development objectives. The customary criteria for evaluating scientific research, most notably journal publications and instances of citation, are not comprehensive enough to consider the appropriateness of technology or its value to development. Research organizations must, therefore, broaden the classical criteria for evaluating research, and monitor these more comprehensive indicators. Some organizations are trying to evaluate user satisfaction as the major criterion of ex post evaluation. This can be difficult, as farmers may not be able to place a value on the research itself.

Advance preparations for ex post evaluation should include precise plans on documentation to obtain, people to interview, and sites to visit. Some supplemental information may need to be gathered through surveys or interviews. Most evaluation teams use a blend of interviews, field visits, observations, and report reading. Beyond a comparison of achieved and expected results, ex post evaluation tries to clarify the internal and external factors affecting these results.

Impact Evaluation

Impact evaluation, also referred to in some of the literature as ex post evaluation, attempts to determine the extent to which research programs have contributed to larger development goals, such as increased farm production or food self-sufficiency. Impact evaluations, which often indicate rates of return on the research investment, are primarily used to convince policymakers to allocate more resources to research. It has a time frame of ten or more years after research results have been released, making it less of a management tool than the other types of evaluation. As with other ex post evaluations, the baseline data, targets, and assumptions from ex ante evaluation are the basis for determining progress and ultimate impact.

An evaluation of the impact of research programs on development is greatly facilitated by prior evaluations of each project. If project and program evaluations are eventually to be used to support impact evaluations, this should be considered during ex ante evaluation and indicators chosen to serve this future purpose. For instance, if market prices need to be monitored periodically for use in a future impact evaluation, this must be identified at the ex ante stage of evaluation and monitored during the course of the activity.

Evaluating the contribution of research to economic development is complex, and there has been controversy over some of the techniques

used. Impact evaluations must distinguish between the contributions research makes to national development from the contributions made by other factors, such as the existence of a good extension service, agricultural inputs, adequate infrastructure, and favorable marketing and pricing policies. Whereas benefits have been relatively easy to attribute in the case of single-commodity technologies, such as high-yielding rice under irrigation in Asia, it has proved far more difficult to do this in more diverse and complex agricultural systems. As a consequence, input-output models which simply correlate investment levels in research with national productivity are not convincing.

Therefore, in addition to economic analysis, the case study approach has often been adopted, seeking the views of the users of the research where they can be identified. While this approach can give a more complete picture of the contribution of research to the client community and ultimately to national development, this approach requires extensive and often expensive data collection and a thorough analysis of socioeconomic factors. The results of impact evaluations can have broad implications for future priority setting, not only for research but also for development support services.

EVALUATION CRITERIA AND METHODS

It is useful in a discussion of monitoring and evaluation to distinguish between the terms criteria, indicators, methods, and techniques. Research evaluations are undertaken for many reasons and in many different contexts, and the evaluation criteria under investigation vary to some extent, depending on the type of research under investigation (e.g., basic, applied), the time frame of the evaluation (e.g., ex ante, ex post), and the aspects considered important (performance, quality, and relevance). Indicators are then selected to serve as a measurable means of determining whether the criteria are being met. They are usually quantifiable and can be verified. For instance, if a criterion for project selection (ex ante) is scientific merit, the indicator used might be demonstrated usefulness of the technology to an associated ongoing program, and the method used might be a form of peer review. The diversity of possible choices of criteria and of the methods used to evaluate them, makes the development of generalities difficult. For any given method, there might be several techniques, the use of which will depend on available data and personnel to conduct the analyses. This paper concentrates on general evaluation criteria and methods, and not on the specific indicators and techniques employed.

For project selection (ex ante evaluation at the project level) the criteria generally applied are technical and operational: scientific merit, the appropriateness of the resources required, and whether the project addresses larger program objectives. The need to demonstrate potential usefulness of research to greater social and economic goals is clearly strongest in mission-oriented research, but in recent years this criterion has also become important for strategic research. Strategic research targets are more difficult to define, and evaluation of strategic research tends to be more continuous and less product-oriented.

Ex ante evaluation criteria at the program level are expanded to accommodate an increasing emphasis on relevance and national returns on investment, and include scientific merit, relative opportunities for success from major thrusts within a commodity or among commodities, the allocation of resources to cover major program objectives, and the relationship between program objectives and national research objectives.

The criteria for ex post evaluation are related to those used in ex ante evaluation. However, they are most often conducted for large projects, programs, or institutes, and are not considered cost effective for all individual projects. The criterion of scientific merit remains important; this can be hard to assess because the component projects can be difficult to aggregate and analyze. Ex post evaluation criteria may differ from those originally defined in ex ante evaluation. Economic and social impact of a program may have become more important than originally assumed, or the objectives of the institute may have changed to reflect changes in policy or opportunities for development. Ex post evaluation may also be more concerned with efficiency. Evaluation criteria and what constitutes success usually differ from one project to another; this means that comparison with other projects is limited as an evaluation method.

Several general methods are used in research evaluation, depending on the criteria considered most important and the purpose of the evaluation. For instance, if performance is the primary concern and the purpose of the evaluation is to improve institute or program management, resources and processes will be monitored and evaluated in what is often called a "performance audit". If research quality is the primary concern, peer or expert review in some form will predominate. For relevance issues, the primary method is comprehensive evaluation based on technical and socioeconomic analyses, using experts from various disciplines. It is very important that those carrying out the evaluation and those being evaluated understand its purpose, and the intended use of the results.

Techniques refer to different tools used to supply evaluation teams with information, and these techniques are being continually refined and developed. It is important to note that while there is controversy over certain techniques, there is at present little disagreement about the fundamental soundness of methods of evaluating research based on a consensus of experts.

Currently, the predominant method employed in the evaluation of all three aspects of research is based to some degree on peer or expert review. Peer review is the process in which scientific merit is evaluated by other scientists working in or close to the field. The field must be sufficiently large for peers to be readily identified and there must have been time for the field to develop to the stage where a basis for agreement on what constitutes quality exists. By contrast, expert review involves the eminent specialists in a particular field, and is more often used when evaluating an entire program or a specialized institute.

There are numerous ways of conducting peer and expert review, and the method selected must take into consideration the organization and the culture it operates within. While some research organizations have had

negative experiences with peer review, there is no substitute for it in research evaluation. In these cases, the organization should attempt some modified peer review process. The Organisation for Economic Co-operation and Development (OECD) has identified the following types of peer review:

- Direct peer review is defined as a review by scientific peers which is confined to determining the scientific merit of an activity. Committee peer reviews are common, both for ex ante and ex post evaluation. Committee members may reach decisions individually, through a group consensus or in a phased combination of the two. Criticisms of the method arise when it is applied to multi-disciplinary research, or in cases where resources are severely constrained and peers may be in competition with those being reviewed.

- Modified peer review is similar, but the criteria are broadened from scientific merit to cover the socioeconomic aspects of strategic and applied research. This requires integrating non-scientists into the direct peer review process. The most frequently used approach is to include users of research on committees and panels. A two-stage process is often used: one which looks at "good science" and the second which looks at relevance. Another approach is to supplement the conventional direct peer review with interviews and/or questionnaires to add more and different information to the evaluation. Many evaluators have suggested that this method is less reliable than open committee discussion in obtaining a balanced view. In any event, balanced judgements are necessary between specialists and generalists, and the management of the process becomes extremely important.

- Indirect peer review is based on information from peer reviews conducted for other purposes. Distinction in scientific performance is recognized in many ways: through the award of prizes, membership in prestigious societies, honorary degrees, etc. Other indications of scientific performance come from bibliometric analysis. Bibliometrics is founded upon the assumptions that the output of scientific research is consistently represented by articles appearing in scientific journals; that the number of citations to these articles is a legitimate indicator of their impact or quality; and that accurate data are available. It is questionable whether bibliometrics is relevant in developing countries. Some problems include the time and expense of conducting the analysis, a bias towards English-language journals, poor coverage in the data base of certain fields, and in applied and adaptive research.

There are many problems associated with peer review. Successful peer review depends on evaluator objectivity, true scientific expertise, and a common objective of improving research. In fact, this is often not achieved. Depending on the personalities, skills, conflicts, and competitiveness within the organization, peer review may be negatively applied. The inclusion of foreign experts can provide a means of achieving objectivity in peer review where competitiveness or lack of scientific expertise exists. The problem of scientific objectivity also becomes greater when the research is multidisciplinary or aims to achieve social or economic objectives. It then becomes necessary to broaden the

team of evaluators to include specialists from different disciplines, and perhaps representatives from development organizations, planning bodies, etc. The institute may need to develop team building and communication skills among the evaluators to do this effectively.

Attempts are being made to enhance peer review by introducing a broader base of information, such as through bibliometrics, but in no cases have these methods supplanted the need for peer or expert review of some kind. These methods, which rely heavily on publications and citation indices, are also considered less relevant to developing countries conducting applied and adaptive research. More significant is the growing use of methods such as questionnaires and structured interviews to gather information. These can reach large numbers of practicing scientists, development workers, extension services, etc., thus bringing more information to the traditional peer review process.

SETTING UP A MONITORING AND EVALUATION SYSTEM

What are your M/E requirements?

An M/E system is not an end in itself. It is successful only if it can be used to improve research. It costs money and time, and so each national research system or institute should determine its own requirements for M/E, and develop a system which responds to these needs. M/E are used at different levels of management for different purposes. In general, three important aspects of research are to be evaluated: performance, quality, and relevance. The following questions may serve as a guide to research organizations considering setting up or enlarging their M/E systems:

- Which aspects of research are weakest in the organization - performance monitoring, quality control, or relevance of research? The performance of research deals with the efficient use of resources and the achievement of stated outputs on schedule. The quality of research is concerned with the adherence of researchers to accepted standards of scientific work. The relevance of research has to do with the appropriateness of the original objectives and the potential of research to address important national objectives. Monitoring and evaluation have a role to play in all three aspects.

- Is research planning at all levels sufficient to provide clear objectives and targets? The relevance of research can only be determined if its objectives are clearly stated and relate to greater development objectives. Likewise, evaluation cannot check performance against objectives if these objectives are poorly defined. At each level of research management, objectives must be explicit. For instance, for long-term planning the stated objectives might include the preparation of a manpower training plan and an infrastructure development plan within certain defined budgetary limits, and the identification of researchable problems with targets in major commodities for the next 10 years. The need for a clear specification of shorter-term objectives and targets is also needed at the research project level, where both performance and quality are important elements.

- Are new M/E procedures necessary? Is the activity already assessed for performance, quality, and/or relevance? What are the shortcomings of the current process? Is a more formal approach needed? Sometimes the M/E system is not the problem; management may receive adequate information to apply remedial action, yet fail to act for lack of management skill, resources, authority, etc. In such cases, nothing is gained by installing a more elaborate M/E system.

- Who needs the information, for what purposes, and in what time frame? It is essential when planning M/E activities to clearly determine who will use the results and for what purposes. The level of effort of M/E should be scaled to the expected use. Also, unless data gathering, analysis, reporting of information, and subsequent action are going to be timely, there is little point in setting up an M/E system at all. To assure this timeliness often requires that authority for corrective action be delegated to middle-level managers who are closest to the problem.

- What indicators are needed to do the job? Once objectives have been defined, verifiable indicators of achievement must be selected, and the methods of measurement determined. These provide the basis for monitoring, and systematic monitoring builds up the body of information for subsequent evaluation. The choice of indicators has implications on the type of people needed to monitor them, and the costs of M/E. While one might be tempted to assemble volumes of data, especially if a computer system has recently been installed, data are costly to collect, and irrelevant data will only complicate analysis later on.

- Are the data to be collected objective and verifiable? Objective in the sense that the same data collected by different people would be consistent; verifiable in that recorded data exist to back-up any conclusions made.

- Are personnel and funds available to do the M/E work? A commitment of resources is necessary at all levels to record and analyze data, report information, and carry out evaluations.

Is a separate M/E unit needed?

Many research organizations, having once made the commitment to do more thorough evaluations, decide to establish a separate monitoring and evaluation unit. The evaluation function is often combined with a responsibility for programming future activities, in a planning and evaluation unit. Whether a research organization needs a special unit for monitoring and evaluation depends on the types of evaluations deemed necessary, and the quantity and type of data required. For example, extensive and specialized data may be needed to perform detailed economic analyses, notably for impact evaluation. The complexity of the research organization - its size and structure - will also influence whether or not a special M/E unit is needed.

If a special unit is created, its staffing will obviously depend upon the functions assigned to it. Much of the information needed to monitor and evaluate research is obtained by research managers themselves, through

regular reporting. Unless very extensive data collection and analysis are undertaken, a specialist in evaluation should not be necessary. Some expertise, however, may be required on a consultancy basis to set up the unit. The unit should have a service orientation, and be staffed, at least in part, by former researchers. M/E will be better accepted by researchers if it is perceived as an internal activity, conducted by fellow researchers who are sympathetic to the special nature of research.

How much does M/E cost?

The costs of monitoring and evaluation depend upon the services required, the extensiveness of data collection and analysis, and whether external assistance is used. While much of M/E is done routinely by researchers, costs must be budgeted for the direct costs of the evaluation staff, for travel, for coordinating and tabulating data at the program and directorate level, and for preparing evaluation reports. There is also a less visible opportunity cost of the experts' time, and the costs of disruption in the organization. Though costs vary considerably from one organization to the other, a range of 0.5 to 1.0% of the program budget is often cited for direct costs. These costs should be included as a separate budget line item for all major research programs.

The costs of evaluation, particularly of ex ante evaluation, should be balanced against the benefits to be gained from sound research planning. Evaluations can provide less obvious benefits also, such as an improved channel of communication within the organization, or an unexpected exchange of technical information. This is particularly so when evaluations are multi-disciplinary. In summary, while costs may vary, they should be proportionate to the importance of the program or institute being evaluated.

Guidelines for evaluations.

Regardless of the type of evaluation performed, general guidelines are available to make the process more efficient and effective. First, the objectives and boundaries of the evaluation must be outlined in advance, and should be agreed by both the evaluators and those being evaluated. For an annual review this is not complicated, but for more in-depth evaluations, a "scope of work" statement may be necessary. This statement lists the objectives of the evaluation, and the type of information to be obtained and analyzed during the evaluation It is a statement of expected output, of questions to be answered. It does not usually specify the precise methods to be used.

The plans for a comprehensive evaluation indicate in advance the documentation needed, people to interview, and sites to visit. It will identify the supplemental information which may need to be gathered through surveys or interviews. Most evaluation teams use a blend of interviews, field visits, observations, and report reading. As research plays a role in national development, an evaluation must consider the institutional, political, social, and economic context in which it is conducted. Beyond a comparison of achieved and expected results, an evaluation should try to clarify the internal and external factors affecting these results.

For comprehensive program or institute evaluations it may be desirable to bring in expertise from outside of the organization. The objectivity and specialist skills of external evaluators can be an asset, especially for more complex or controversial research, for reviews of programs within institutes and their complementarity, and for suggesting major changes in the organization's thrusts. The major disadvantage to external participation is the cost and a lack of familiarity with the country or the organization (Table 2). Regardless of the size and cost of the effort, evaluation should be considered part of research planning and implementation, and therefore incorporated as a regular research expense.

Table 2. Considerations for Using In-house Versus External Evaluators

In-house Evaluators

Advantages	Disadvantages
* Familiarity with programs & staff operations	* Objectivity and candor may be questioned
* Consistency assumed with national value system	* Possibility of organizational role conflict
* Less time required to schedule evaluations	* Difficulty of releasing staff from regular duties
* Less expensive	

External Evaluators

* Greater objectivity	* May be perceived as policeman and make staff anxious
* Free of organizational bias	* Requires time for contract negotiations and orientation
* Possibly greater access to decision-makers	* More expensive
* Time exclusively devoted to task	
* Familiar with recent advances in technology	

Collaborative Evaluators

* Advantages of both in-house and external evaluators; plus broader cultural and technical perspective	* Some candid discussion of sensitive national issues may be constrained

Modified: USAID. Design and Evaluation of AID-Assisted Projects, 1980.

An evaluation culminates in recommendations to the appropriate level of management: for a research program this might be suggestions for the revision of its objectives, workplan and schedule, or the suggested termination of a project. At a higher level a recommendation might be made to provide emergency funds, change staff assignments and priorities, or radically modify or terminate a program. Those researchers most involved in the activity being evaluated should be given the opportunity to comment on the evaluation and the recommendations made. For any evaluation to be useful, feedback and corrective actions must be timely.

ISNAR'S APPROACH TO DEVELOPING M/E MATERIALS

Because of the diverse contributions monitoring and evaluation can make to research management, ISNAR cannot develop M/E materials in all areas at once, this all the more so because of the dearth of background materials from which to draw. One annex on monitoring and evaluating ongoing research activities has been produced thus far. The primary focus is on evaluating the effectiveness and efficiency of the research program, including decision-making procedures, program leadership, and the adherence of scientists to accepted standards of research execution and reporting. Important associated elements of personnel management are presented in the ISNAR paper, "Human Resource Management in Agricultural Research".

Planning or ex ante evaluation is probably the single most important activity of a research manager. ISNAR has prepared a separate paper, "Priority Setting in Agricultural Research" which describes different planning techniques at the national level, including simple checklists, scoring models, and benefit/cost analysis. A more thorough paper on ex ante evaluation at the program and project levels is being produced; it will provide more information on establishing baselines, defining targets, and identifying criteria and indicators for subsequent monitoring and evaluation.

An additional annex will eventually be written on ex post evaluation, the measurement of performance, quality, and relevance of a completed research project or program. Ex post evaluation provides the basis for incorporating lessons learned into subsequent planning exercises. This evaluation uses the information from ex ante evaluation as its reference if it exists.

The development of materials for impact evaluation of research programs has not yet been undertaken by ISNAR. Impact evaluation of research is an object of study in many universities and other organizations. It is an evaluation whose greatest contribution is to policy dialogue, but it has limited use as a management device. At the present time, ISNAR's resources are being used to develop approaches and methods in ex ante, ongoing, and ex post evaluation, which are more useful to research managers in the medium term.

Annex 1

MONITORING AND EVALUATING ONGOING RESEARCH

Definitions of terms and concepts

The terms "program" and "project" are used in various organizations to refer to different things. In this paper, the term program refers to coordinated research activities whose combined scientific output addresses national research objectives. Programs are long-term in scope, and somewhat continuous. They are composed, in some cases, of sub-programs, and of projects. Projects address specific research problems, and have explicitly defined time frames, resources, and targets. Each project in turn comprises a number of specific operations or experiments.

For example, a maize program might be created to address the research objective of improved maize technology, which itself responds to development objectives of increased cereal production. It is divided into two main sub-programs: maize breeding and agronomy. Within the agronomy sub-program there are three projects: maize fertilization, maize intercropping, and socioeconomic study of local practices. Operations within the maize fertilization project might include trials using rock phosphates, compound fertilizers, and manuring.

The complexity of a nation's agriculture and the size and structure of the research system will determine the appropriate hierarchy of programs, sub-programs, projects, and other subdivisions. However they are defined, objectives must be determined for each level of the research hierarchy. In turn, research evaluation should take place at all levels in reference to these objectives: from experiments and operations, to individual projects and programs, to programs in the aggregate. This ensures the soundness of technical decisions from program planning to the execution of experiments. By routinely evaluating research activities, researchers and supervisors can modify these activities in the course of program implementation in light of information gained.

Relationship between ex ante and ongoing research evaluation

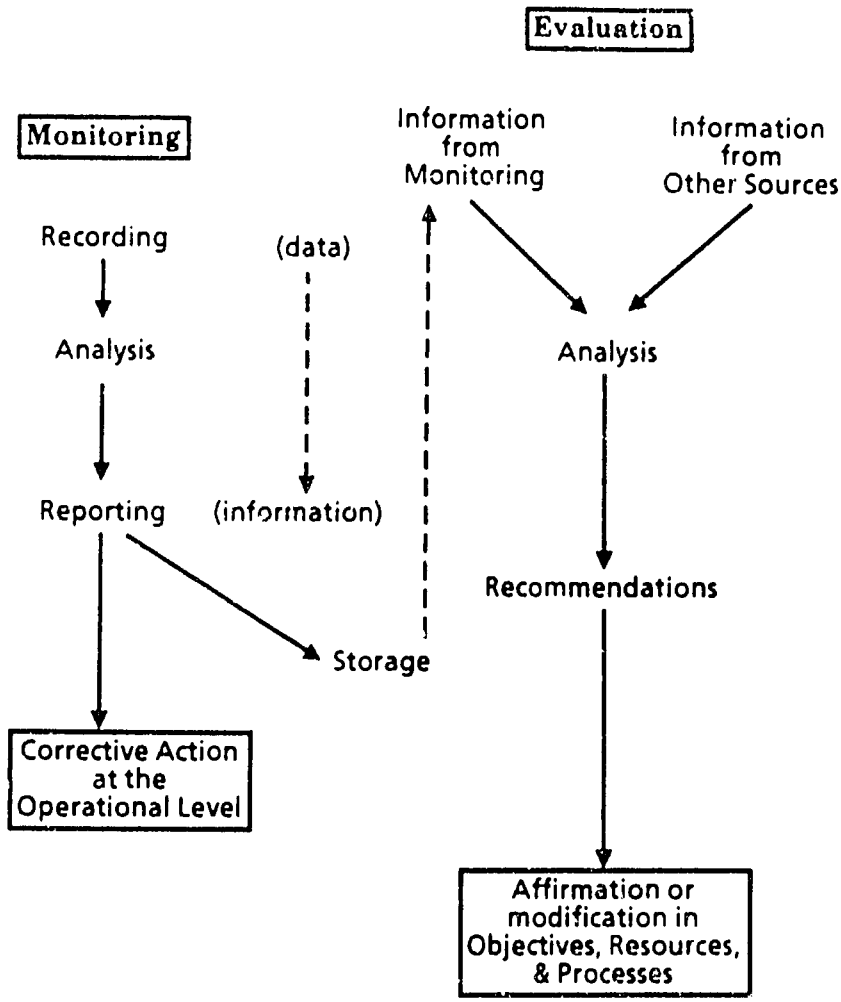
Good planning or ex ante evaluation is necessary for the evaluation of ongoing research. It is against the original plan - its targets and assumptions - that an activity is judged. There are, of course, different levels of ex ante evaluation in national agricultural research. The highest level determines how well the research system responds to national development objectives. It can involve priority setting and broad resource allocations across commodities, regions, and disciplines. At the subsequent echelon of decision making, resources are allocated to alternative research thrusts or approaches within commodities, based on an analysis of the technical constraints and opportunities, and the availability of resources. At the next level, programs and projects within programs are defined and alternative research activities selected.

At every level of program planning there should be clear definitions of objectives, identifiable inputs and expected outputs, and some notion of time frame. Any assumptions used in planning, whether for a program strategy or a project proposal, should be explicit and may be subject to evaluation later on. It is during ex ante evaluation exercises that the key indicators used to monitor and evaluate progress are identified.

Monitoring ongoing research

Monitoring is primarily focused on research project performance, the use of resources and the achievement of outputs. It comprises several processes: the regular recording of key indicators; analysis which transforms data into information; reporting the information to appropriate levels of management; and a system of information storage which is easily accessible (Figure 1). These processes are integrated into day-to-day management for the purpose of keeping an activity moving as planned. Monitoring is primarily focused on inputs and outputs, on whether personnel, buildings, equipment supplies, funds, and support services are available and on schedule.

Figure 1: Relationship of Monitoring to Evaluation



Data are available on research activities through various sources. Administrative personnel may keep detailed records of staff time, funds, and equipment used. Technical data are recorded in field notebooks, periodic research reports, and annual project reports. If these data are systematically collected and analyzed in a timely way, problems can be identified and resolved, and information valuable to other researchers can be circulated. While most data used in monitoring are recorded by researchers or administrative staff, project leaders are usually responsible for their compilation and analysis. There should be clear assignments of responsibility for data collection, analysis, reporting, and remedial action.

Monitoring and evaluation are facilitated if a project management approach is used. The key indicators to be monitored are identified at the outset when drawing up the research project workplan. A Logical Framework matrix, described in ISNAR Working Paper No. 12, is one technique that can be used for defining indicators and sources of data (Table 3). The link between what is scheduled and what actually takes place can be directly established by comparing periodic progress reports with the original workplan. This not only forces project leaders to keep track of their activities, it also facilitates communication of problems to higher levels of management.

In the instituting of a new monitoring procedure, it is necessary to start with an analysis of the current situation: what reporting is currently required, how well is it accomplished, and what additional reporting is necessary? The functions of a monitoring system should be agreed upon by the people implementing it. Several questions should be answered when designing a monitoring system:

- Who needs the information and for what purpose?

- What are the simplest means possible of collecting the necessary data? Can they be obtained from existing sources? If not available, can they be collected at reasonable cost in relation to their usefulness?

- Can the information resulting from data analysis be presented in a simple, standard format for timely use in decision making?

- Can the information be stored in a format compatible with that from other sources, so findings from similar activities can be compared?

With the increased use of microcomputers in developing countries, it is appropriate to determine which monitoring procedures, if any, should be computerized. Microcomputers can be very useful for analyzing and storing large volumes of data, but they are not necessary for all research organizations or for all M/E processes. A lack of computer capability should never be accepted as an excuse for inadequate record keeping, just as the existence of computers does not justify increasing the quantity of data gathered. On the other hand, once a good data base has been created, it becomes easy with computers to update it periodically and to retrieve data in various forms.

Computers can be especially useful for maintaining a data base on personnel, for program budgeting in larger research systems, and for management of physical and financial resources. Computers certainly facilitate the

Table 3: Examples of Research Program Indicators

Levels of management	Possible indicator	Means of verification	Responsibility for data collection
Project Inputs - based on operation workplans:			
- personnel	scientific and support time	time sheets	individual reports
- funding	expenditures	accounting data	accounting office
- facilities	construction or acquisition	on-site report procurement data	institute engineer accounting office
- equipment & supplies	acquisition actual use	procurement data lab/station logs	accounting office lab/station manager
- leadership	project meetings program meetings	meeting reports meeting reports	project head program head
- training	courses completed	training records	training officer
Project Outputs			
- preliminary research results	research data from experiments	research reports publications and surveys	scientist project head
- completed research results	program committee recommendations	program records annual reports	program head NARS director
- research capacity improved	trained personnel & improved facilities	training records administration records	training officer administrator
Program - Contribution of knowledge from research programs to research, development and policy-making bodies:			
- new knowledge of interest to research, extension & policymakers	released technology or recommendations	program records certification res/extension communications on policy	program head national body extension service NARS director
National/Institute - Research relationship to national development objectives:			
- increased crop production	production data	farm surveys	statistics dept.
- intensified land use	changes in crop patterns & inputs	input statistics	devel. ministry
- conservation & land use	reduced erosion resource planning	survey methods planning document	land use body planning body
- increased income	per capita change increased spending	national data village surveys	statistics dept. devel. ministry
- improved nutrition	decreased disease & mortality	nutrition surveys	national health service

use of management tools, such as the critical path analysis described in ISNAR Working Paper No. 13, when scheduling tasks and tracking progress. Data collection systems should be designed taking into account whether techniques will be done manually or on computer; if computers are to be used there can be advantages to decentralizing the inputting of data on microcomputers at research stations. This minimizes the risk of transporting data and makes management at these stations responsible for the accuracy of these data. It is also important to select the software for tabulation and analysis before designing the forms on which data will be recorded.

The most common form for reporting information in monitoring is the periodic progress report. Researchers and project leaders are often required to submit such reports on resources used, problems encountered, and tasks achieved. They signal management to take remedial actions. Progress reports can be used in conjunction with program budgeting forms and project workplans to check progress against planned objectives. Progress reports do not necessarily require a narrative text, and they can be further simplified by using a standard format which refers back to the original workplan.

A simple format for reporting facilitates analysis and synthesis of information. For instance, program leaders can much more easily review the project progress reports and summarize them if the format is standard and problems in implementation are easily identifiable. In cases where related information is reported from different sources, a similar format permits easier cross analysis. At each level of management (project, program, institute) some degree of reporting is required (Table 4), and if information is not summarized at each level as it moves upward, the directorate will receive unmanageable quantities of information, often of unnecessary detail.

Storage is another important aspect of a monitoring system. Whether manual or computerized, information must be accessible. A centralized system of information storage can cause unnecessary delays at outlying stations that need it. Conversely, datailed back-up data from stations are not necessary in central files. When deciding where to store data and information, the need for it, the facility of movement, and the security of storage should all be considered.

On a purely technical level, research managers also monitor the potential success and acceptance of a technology during the course of its development. This can be done through the early introduction of on-farm research to study socioeconomic aspects of the proposed technology and to verify whether technology being developed is appropriate. Several techniques are available to gain information on farmers' constraints, interests, and capabilities through this process. Other techniques for getting feedback on adaptive research are formal and informal surveys of attitudes and practices. All of these feed information into ongoing research and can be thought of as monitoring techniques.

Table 4. Flow of Monitoring Reports

Person responsible	Report	Sent to whom
Accountant	Monthly or quarterly budgetary statement	Researcher/Project leader Chief accountant
Researcher	Periodic progress report includes financial summary	Project leader Station head
	Annual/final research report	Project leader Station head
Project leader	Summary of progress reports	Program leader
	Annual project summary	Program leader
Program leader	Summary of progress reports	Research director
	Annual program meeting minutes and report	Research director
Station head	Summary of station data on resources used and needed	Directorate officers in charge of funds, supplies, and facilities
Research director	Annual report	Technical, planning and finance ministries
		Donors

Evaluating ongoing research

Evaluating ongoing research, also called interim evaluation, looks beyond the monitoring of performance, and also considers aspects of quality and relevance, the latter being largely achieved through careful ex ante evaluation. While evaluation is an important management tool, it should be underscored that no M/E system can make up for a lack of key scientific leadership, as so often occurs in research organizations where staff are young and inexperienced. There are, however, some mechanisms, such as the rigorous review of project proposals, regular reporting, and an annual defense of research activities, that help those existing senior scientists supervise the work of younger scientists. Evaluations of

ongoing research take place annually in the case of projects, and at greater intervals, say every 3-5 years, in the case of entire programs. The essential elements of these different types of evaluations will be discussed here, commencing with a brief introduction to the importance of peer review in the process.

Role of peer and expert review. Research evaluation examines aspects of performance, quality, and relevance. All aspects imply an assessment of the technical judgement and skill employed in the research, and the organizational support and resources provided. Peer review, the assessment of research by colleagues, is the approach often taken in annual project reviews; expert review is more often applied to comprehensive program reviews. Peer and expert review work best where constructive criticism and open discussion are acceptable.

Peer review can play a role in many aspects of research, such as whether project and program objectives continue to address important research priorities; the degree to which planned objectives have been achieved; the need to modify objectives or to propose activities in light of new information, the quality of the research conducted; its efficiency relative to other projects; the timeliness and effectiveness of research communications; and the need to develop interdisciplinary teams for implementation and analysis.

Role of strong program leadership. In many developing countries the majority of scientists are young and relatively inexperienced when they assume research responsibilities. For this reason, many research organizations in developing countries have adopted a project approach to research, versus a researcher-oriented one allowing maximum autonomy. Strong program leadership and supervision of less-experienced staff are indispensable in these research systems, though care must be taken not to stifle initiative or innovative thinking. The project management approach adds transparancy to the research organization, and requires a certain rigor in planning; this lends itself to involving more-experienced staff in the process of proposal development.

Annual project and program evaluations. Good program evaluation is based on an annual review of component projects. The annual project evaluation is a peer review of completed and proposed experiments, and includes a discussion of objectives, rationale, and methodologies. Senior scientists play a crucial role in these reviews, structuring the discussions and suggesting alternative ideas and approaches to less-experienced researchers. Senior research management should be able to assess the quality of this process if good minutes are written up of the meeting, including the topics discussed, issues raised, and recommendations made.

The annual project review depends on the timely submission of annual project reports. Writing these reports seems to be a bottleneck in many NARS. Preparation for and participation in these technical reviews, including prompt report writing, is part of a researcher's job and, therefore, provides one possible criterion for employee performance appraisal, and the application of rewards and sanctions.

Annual reporting usually includes a summary of the objectives of the project, methods, data analysis, results, and discussion. The length and complexity of this annual report will depend on the type of experiment and the state of advancement of the project. A summary account of resources used and costs are attached, based upon the periodic progress reports and the original workplan. It would be beneficial if a final research report on a completed project were more comprehensive, taking the form of a scientific report. This would include an abstract, research objectives, materials and methods, results and discussion, conclusions, and references. Should the scientist wish to publish in a professional journal, this report need only be edited and modified to conform to particular journal requirements.

Annual program evaluation and planning meetings require the active participation of researchers, project leaders, and program heads. Experts from outside the institute may be invited to participate. These reviews deal with the scientific/technical content and resource requirements of all projects which comprise the program. They also provide a forum for discussing problems in program management and certain aspects of station management. The annual program evaluation and planning meeting is monitored through minutes, and an annual program report summarizes the year's activities and the decisions taken to initiate, continue, strengthen, or terminate projects. The institute director and/or research committee then develops the plan and the proposed budget of the institute for the following year based on reports and recommendations of the annual program reviews.

Comprehensive program evaluation. Periodically, say every three to five years, it is important to organize program-level evaluations that are more in-depth than those occurring annually. These more comprehensive program reviews are both formative, that is, results from the evaluation are used to modify and improve the existing program, and summative, where decisions are taken to initiate, continue, or terminate programs. The emphasis in program evaluation tends to be on current and future programs, using lessons learned from past research in a forward-looking way. The original objectives, targets, and assumptions of the program are analyzed to see if they are still valid. They are modified accordingly. Project approaches and their likelihood of meeting existing or revised program objectives are examined. They are also viewed in the aggregate, in terms of their importance, resource requirements, and complementarity. This level of program analysis is necessary for setting priorities, allocating resources, and defining long-term research plans at the national and institute levels, as well as for making effective use of program funds. Comprehensive program reviews require experienced scientific leadership, and are greatly facilitated if good annual project reviews have taken place and are well documented.

Many of the available methods and techniques of program evaluation are currently more often used in disciplines other than agricultural research (Table 5). The evaluation of research requires some modification of existing techniques to accommodate the uncertain nature and timing of research, and the interdependence of different technical disciplines. This topic has come under considerable study in recent years, but techniques for research program evaluation are not yet fully defined.

Many different terms are used in the literature to describe the methods and techniques used in program evaluation. The central features of all these approaches are that they are analyses of program processes, not just program content. They, therefore, have implications for improving efficiency and effectiveness. They include quantitative and qualitative techniques. The primary method in program evaluation is informed judgement through peer or expert review. This body is supplied with information on programs through interviews, observation, and analysis of background data. Modeling, simulation, statistical analysis and other quantification techniques are often used. In some cases, analyses of qualitative data using the analytical and data collection methods of social scientists may be appropriate.

Table 5. Techniques Used in Program Evaluation

Checklists	Consultation with Experts
Content Analysis	Cost-Benefit Analysis
Cost Effectiveness Analysis	Cross-Impact/Cross-Support Matrix
Decision Analysis	Delphi
Input-Output Analysis	Interviews and Opinion Surveys
Historical Analog	Least Cost
Modeling	Network Analysis
Probability Tree	Public Participation
Rank Size Analysis	Relevance Tree
Scoring Models	Simulation
Statistical Analysis	Stochastic Estimates
Substitution Curves	Trend Analysis
Workshops, Panels, Conferences	

Modified: Coates, Vary T. "Technology Assessment in Federal Agencies: 1971-1976". National Science Foundation, March 1979.

Most research programs in developing countries are responsive to larger development objectives. A comprehensive program evaluation should include, therefore, representatives from development and extension organizations, and a mechanism for bringing user feedback into the process. Program evaluations may also include representatives from planning and finance ministries, depending on the size and importance of the program. Evaluators, particularly for multidisciplinary research which has social and economic dimensions, will always bring with them certain biases. These can be reduced by using standardized checklists, by training evaluators in advance in evaluation methods, team-building techniques, and communication, and by using external evaluators, when possible.

Summary

In summary, this annex has tried to illustrate the importance of integrating monitoring and evaluation activities into day-to-day management practices in national research organizations. It concentrates on the internal monitoring and evaluation which should take place for ongoing research, and focuses primarily on the necessary reporting requirements of the researchers themselves.

On an annual basis, project and program management are most concerned with performance monitoring, and the quality of the research being conducted. The scope of comprehensive program evaluation is broadened, however, to more thoroughly consider aspects of the relevance of research. The primary method employed for all ongoing research evaluation is peer or expert review. There is currently much interest in developing and refining techniques, some borrowed from the social sciences, to make research program evaluations more informative.

Two other ISNAR working papers have been produced which are relevant to ongoing research evaluation. The Logical Framework in Research Planning and Evaluation, Working Paper No. 12, relates project-level activities to larger program and national research objectives, while at the same time identifying targets and assumptions upon which the success of the project and future evaluation are based. Project Management Techniques for Performance Monitoring, Working Paper No. 13, explains the uses of Task Lists, Bar Charts, Milestone Charts, and the Critical Path Network for tracking resources, outputs, and time. Both of these working papers are available from ISNAR for use as companion texts to this paper.

REFERENCES

Agency for International Development. Design and Evaluation of AID-Assisted Projects, Training and Development Division, AID, Washington, D.C. 1980.

Brazzel, John M. "Evaluation of Agricultural Research: Program Evaluation Methods and Approaches", Evaluation of Agricultural Research, University of Minnesota. 1981.

Casley, D., and D. Lury. Monitoring and Evaluation of Agriculture and Rural Development Projects, Johns Hopkins University Press. 1986.

Casley, D., and K. Kumar. Project Monitoring and Evaluation in Agriculture, draft, World Bank, Washington, D.C. 1986.

Coates, Vary T. Technology Assessment in Federal Agencies: 1971-1976. National Science Foundation. March 1979.

Dagg, Matthew. "Working Group on Monitoring and Evaluation of Agricultural Research", Internal Program Review background paper. ISNAR. 1986.

Daniels, Douglas, editor. Evaluation in National Agricultural Research, International Development Research Centre, Ottawa, Canada. 1987.

Hariri, Ghazi. "Monitoring and Evaluation of Research Programs as a Tool for Management", paper for Damascus workshop, ISNAR. 1987.

Financial and Administrative Management of Research Projects in Eastern and Southern Africa (FAMESA). Management Manual for Productive R&D: Strategic and Project Planning, and Budgeting. International Centre of Insect Physiology and Ecology, Nairobi, Kenya. 1984.

International Service for National Agricultural Research. "The Management of Agricultural Research", seminar, Mananga workshop. 1984.

Murphy, Josette. "Using Evaluations for Planning and Management: An Introduction." Working Paper No. 2. International Service for National Agricultural Research, The Hague, Netherlands, 1985.

Murphy, Josette. "Monitoring and Evaluation of Agricultural Research", Concepts, Organization, Methods. draft. International Service for National Agricultural Research, 1985.

Organisation for Economic Co-operation and Development. Evaluation of Research: A Selection of Current Practices. Paris, France. 1987.

Practical Concepts, Incorporated. Manager's Guide to Data Collection, Office of Evaluation, Agency for International Development, Washington, D.C. November 1979.