

**SOME PRACTICAL GUIDELINES  
FOR EVALUATION WITHIN  
NATIONAL AGRICULTURAL  
RESEARCH SYSTEMS  
USING THE CHECKLIST APPROACH**

**BARRY NESTEL**



**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 thirteen 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. 22

**SOME PRACTICAL GUIDELINES  
FOR EVALUATION WITHIN  
NATIONAL AGRICULTURAL  
RESEARCH SYSTEMS  
USING THE CHECKLIST APPROACH**

**BARRY NESTEL**

March 1989

***ISNAR***

**International Service for National Agricultural Research**

EVALUATION EXPERIENCESTable of Contents

	<u>Page</u>
FOREWORD	(ii)
<u>Chapter 1</u>	
BACKGROUND	1
1.1 Introduction	1
1.2 Defining Evaluation	2
1.3 Methods of Evaluation	4
1.4 The Clients for Evaluation	8
1.5 The Method of Choice	9
<u>Chapter 2</u>	
WHAT TO EVALUATE	10
2.1 Introduction	10
2.2 Planning, Priority Setting and Programming	12
2.3 The Available Resources	14
2.4 Activities, Achievements and Impact	17
2.5 The Institutional Background	20
<u>Chapter 3</u>	
HOW TO EVALUATE	22
3.1 Defining the Parameters	22
3.2 Who will do it?	22
3.3 How Much will it cost?	24
3.4 When will it be done?	25
3.5 Planning	26
3.6 Coordination	27
3.7 Preparing Background Material	28
3.8 Organizing the Field Visits	29
3.9 Briefing the Team	29
3.10 The Method of Work	30
3.11 Linking to Other Scientists and Potential Users	31
3.12 Assessing Impact	32
<u>Chapter 4</u>	
WRITING AND USING THE REPORT	34
4.1 Planning the Report	34
4.2 The Terms of Reference	35
4.3 Writing the Report	36
4.4 Using the Report	37
<u>BIBLIOGRAPHY</u>	39
<u>LIST OF ANNEXES</u>	40
ANNEX A	Logical Framework Analysis 41
ANNEX B	Indicative Costs for an Evaluation 43
ANNEX C	Examples of Background Tables 45
ANNEX D	Examples of Report Outlines 46

## FOREWORD

This paper has been prepared in response to requests from national agricultural research systems (NARS) managers for assistance in planning and conducting their own evaluations of agricultural research. It is based on the author's own experiences but also draws on writings and comments provided by a number of past and current ISNAR staff, including Anthony Bottomley, Matthew Dagg, Howard Elliott, Peter Goldsworthy, Krishnan Jain and Diana McLean. The author is particularly indebted to Josette Murphy, whose unpublished work at ISNAR has made a significant contribution to Chapter 2 and Annex A.

The paper consists of four chapters and a series of annexes. The first chapter stresses the links between planning, implementation and evaluation and offers a brief overview of current definitions and methodologies used for evaluation. It recognizes the long-term need for the use of a formal benefit-cost approach but postulates that such an approach is currently not feasible in many NARS. A practical, although less rigorous, alternative for initiating the evaluation process lies in the use of the checklist approach, particularly when appropriate economic criteria are used in the selection of checklist questions.

Chapter 2 discusses the preparation and use of checklists and the rationale in creating such lists for various aspects of evaluation. The third chapter examines the operational procedures for conducting an evaluation based on the use of checklists and the final chapter discusses the preparation and use of evaluation reports. In this last chapter particular stress is placed upon the importance of clear and explicit terms of reference for an evaluation team and on the potential role of checklists in addressing this topic.

Since every research institute and program is likely to have both country specific and unique individual features, no attempt has been made to prepare a formal manual or 'cookbook' of evaluation procedures. In preference to this, the paper which follows represents a sharing of ideas and practical experiences. It presents a conceptual approach with respect to checklist-based evaluation and offers some operational guidelines for implementing this approach at the field level.

For those NARS who wish to go beyond the checklist approach and to embark on formal evaluation based on cost-benefit analysis a separate ISNAR publication is in the course of preparation. This will include a discussion on consumer/producer surplus analysis and efficiency pricing as well as both ex ante and ex post benefit-cost analysis and should be of interest to NARS which have the capability to use a quantitative approach.

## Chapter 1

# BACKGROUND

EVALUATION SHOULD BE SEEN PRIMARILY AS A COMPONENT OF A MANAGEMENT INFORMATION SYSTEM WHICH SERVES POLICYMAKERS, RESEARCH MANAGERS, AND SCIENTISTS. FOR MAXIMUM EFFECTIVENESS IT SHOULD BE CLOSELY LINKED TO PROGRAM PLANNING AND MONITORING AND BE CONDUCTED ON A PRE-ARRANGED TIME FRAME (OFTEN QUINQUENNIALLY). THE USE OF FORMAL BENEFIT-COST CRITERIA IN EVALUATION IS DESIRABLE BUT IS CONSTRAINED IN AGRICULTURAL RESEARCH BY THE PROBLEMS OF MEASURING INPUTS (e.g., QUANTIFYING THE VALUE OF BOTH PAST RESEARCH AND ALSO THE INPUTS FROM EDUCATION AND EXTENSION) AND OUTPUTS (e.g., KNOWLEDGE AND IMPACT). FOR MANY DEVELOPING COUNTRIES, NON-FORMAL SUBJECTIVE EVALUATION BASED ON THE USE OF EXPERIENCED PERSONNEL AND CAREFULLY PREPARED CHECKLISTS MAY BE THE MOST PRACTICAL AND COST-EFFECTIVE APPROACH.

### 1.1 Introduction

Agricultural technology management involves the planning, implementation, and evaluation of agricultural research and extension. This paper deals with only one of these closely related topics, namely research evaluation. However, because of the necessary linkages between planning, implementation, and evaluation and between research and extension, it is not possible to examine any of these topics in isolation of the others. Thus an evaluation of agricultural research entails looking at the agricultural technology management system in its broadest sense.

In most developing countries this system lies largely in the public sector, whose agricultural activities usually include a number of development programs or support services. The evaluation of these is normally measured in relation to the physical progress achieved on a planned time frame in terms of meeting defined goals, the number of farmers served, or the area of land receiving specified inputs such as credit or fertilizer. Alternatively, progress may be measured in the form of output, such as tonnes of wheat grown or milk produced.

The use of these types of measurement criteria does, however, present difficulties when they are applied to the evaluation of agricultural research (Dagg 1986). Among the reasons for this are:

Research creates only a potential for change. Its realization depends on many factors outside of research itself. The immediate beneficiaries of agricultural research are development agencies and planners, and not farmers or consumers, although the latter are the ultimate beneficiaries. However, the time lag between the successful completion of research and its impact at the farmer or consumer level may be ten years or longer.

The efficiency of research often depends as much or even more on the skill with which the objectives of the research are selected rather than on the technical skills with which the research is implemented.

Research is intrinsically uncertain in the timing of its progress and needs to be flexible in its approach, depending upon the nature of the results achieved. A lot of good research work yields essentially negative conclusions which are difficult to accommodate in an evaluation framework which is geared to measuring positive progress.

In spite of these problems the rising global concern for accountability by public agencies has resulted in public research organizations being increasingly asked to justify their resource allocation decisions. A number of quantitative techniques designed to approach economic optimization rules have been proposed for this purpose. But, in practice, few quantitative techniques have yet been successfully implemented, possibly because the cost and complexity of doing so may be seen as outweighing the benefits (Shumway 1983). Nevertheless, the pressure from governments and particularly from donor agencies, who provide significant components of agricultural research funding in many developing countries, has meant that national research directors have become very conscious of the need for evaluation. Many of them also recognize its importance in terms of planning and implementation activities and as an integral component of an effective technology management system.

Research directors who wish to establish an evaluation program face the problem of how to structure it so that it: is both practical and realistic in terms of public accountability; meets the information needs of various levels of research management; and is consistent with the personnel and financial resources available to them.

This paper discusses these issues in the context of some experiences that ISNAR staff have had with research evaluations. It addresses the subject in a practical and non-formal manner which should be within the existing capacity of most NARS.

## 1.2 Defining Evaluation

One of the problems relating to the evaluation of agricultural research is the fact that the term 'evaluation' means different things to different people. One of its less fortunate connotations is due to the fact that evaluation activities (including monitoring) have often been established in developing countries as parts of special units associated with donor assistance and have been seen as surveillance activities, rather than as a constructive component of a management system.

The prime role of evaluation is, however, not to find fault but to act as an integral component of an ongoing technology management information system which can help to draw lessons from past experience and incorporate them into the planning and implementation of more effective future activities. Thus, research evaluation is a management tool which can help to set priorities, allocate resources, highlight and clear bottlenecks, identify productive staff and research methodologies and generally help to optimize the return to investment in research.

In this sense the evaluation process should really begin when a program is being designed, through: (a) setting up clear objectives; (b) selecting verifiable indicators of achievement; and (c) specifying how the achievements will be measured. If all this is done it will help force the program planners to clearly express from the outset what are the objectives of the research program and how its results are to be measured. It will

also make it possible to establish a basis for monitoring progress, since it is much easier to conduct a comprehensive evaluation of research if adequate records and reports are available throughout the lifetime of a program. Unfortunately the real situation in many research organizations is such that the technology management system is not well developed. At the time of project planning it is common for rather limited attention to be given to programming, monitoring and evaluation, and verifiable indicators of achievement are seldom available.

Some classification of terminology with respect to monitoring and evaluation may be useful at this point. Many institutions do not sharply define the difference between these terms and the two activities often form components of the work of the same organizational unit. Monitoring is, however, essentially an activity internal to an institution, whereas objectivity is often sought in evaluation by incorporating outsiders into the process.

A United Nations Task Force on rural development has defined monitoring as "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, targeted outputs and other required actions are proceeding according to plan". In contrast to this, evaluation was defined as "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 decisions making (United Nations 1984).

Practitioners of monitoring and evaluation use different terms to frame their discussions. A popular categorization is based on a time framework, and divides monitoring and evaluation into ex ante evaluation, monitoring and ex post evaluation.

Ex ante evaluation is a comprehensive analysis of the potential impact of an activity before it is carried out. It serves to set target objectives and is used in priority setting and resource allocation. It defines the baseline against which progress towards planned objectives would be measured in subsequent evaluations. But in doing this key indicators have often not been predetermined and data may be scanty or unavailable. Nevertheless this type of evaluation, which can lead to improved research planning and the setting of objectives, is becoming a subject of increasing interest to NARS.

Monitoring, which may be regarded as a form of ongoing (or progress) evaluation, involves the collection of data on key indicators and the analysis of this and other information during the implementation of an activity. Comparing achieved with expected results in a given time-frame is the most prevalent and useful form of monitoring for management purposes. It indicates how efficiently resources are being used and identifies problems in management.

Monitoring primarily tracks the use of physical and financial resources towards the stated technical objective. Thus it requires the periodic generation of information on the availability and deployment of staff, infrastructure, equipment, supplies, services and funds. This type of activity is widely used by donor agencies.

It does not, however, measure whether the objectives of a research project or program are being attained and this is carried out more widely through ex post evaluation which assesses performance after a project has been completed. Such an evaluation attempts to measure the efficiency and effectiveness of a completed activity and includes an analysis of constraints and achievements which can be used as lessons learned for future activities.

### 1.3 Methods of Evaluation

A number of different approaches have been used to evaluate agricultural research. Some of these are now regarded as having little to commend them and have been largely discarded. For example, personnel and programs may be evaluated through the number of publications or reports issued, technical meetings held, committees established or seminars undertaken, although none of these criteria necessarily represents tangible activities in terms of bringing about changes in agricultural productivity.

#### Checklists

One of the more widely used methods of evaluation is based on the use of checklists. This is the least sophisticated of the techniques in common use and has the virtue of being the easiest to use. The evaluator uses a list of the criteria and associated questions which he or she wishes to address. The technique is simple to apply but requires considerable understanding of the research and development process.

The checklist approach is sometimes criticised on the grounds that it does not adequately cover economic criteria in a quantifiable manner. However, it must be recognized that there are a number of management issues, particularly those relating to the utilization of the physical and human resources available for research, that do not necessarily readily lend themselves to evaluation by formal methods. In such circumstances the user of checklists by experienced evaluators has much to commend it. The same is true in terms of peer review of biological and economic research.

#### Scoring

Scoring methods represent a more elaborate version of the checklist technique. The scoring matrix is really no more than a checklist with the answers to questions assigned numerical values and weights. Criteria weights are multiplied by the values which a particular research program merits under each criterion to produce a final score. Programs can then be evaluated according to their scores under different criteria.

This approach has some attractions but presents difficulties when evaluating non-commodity research such as soil physics or pest management. It is also difficult to weight maintenance research. However, its main problem is that both the weighting and the scoring process are highly subjective and difficult to standardize across a broad spectrum of research activities.

#### The logical framework analysis

Neither checklists nor scoring provide a quantifiable method of comparing achievements with the targets which a program was designed to meet. To do this requires that indicators of achievement be identified and clearly

stated when the program is being designed. If this is done the information required to both design and evaluate the program can be summarized in a table similar to the Logical Framework that various development agencies use in their project papers (Annex A).

The logframe (McLean 1988a) is a systematic configuration of logical reasoning set out on a four-by-four matrix which helps to organize the various levels of objectives of an activity and sets up some parameters related to the achievement at each level. It offers a methodology which has certain well-defined characteristics and also limitations. It is, for example, neutral in terms of both programming and technology. Unless special provision is made to cover these subjects when the logframe is drawn up it will not, therefore, address what are the most critical constraints nor will it offer guidance on questions of equity.

The logframe can, however, be used as a relating device in that it is fully compatible with a formal and quantifiable cost benefit analysis. At the same time it neither precludes the need for such an analysis nor is it dependent on it.

#### Benefit/cost analysis

Benefit/cost analysis is widely used by governments and funding agencies for deciding on and evaluating investment in development projects. It is based on the concept of discounted cash flow: the premise that a dollar earned a year from now is worth less than a dollar earned today, because of the interest-earning potential lost during the intervening period. Taking into account the effect of time is of major importance, especially in view of the uncertainty of both the time required to complete the research and the subsequent rate of uptake by farmers.

Reasonably accurate estimates of these parameters are crucial for determining the return on resources invested in research. For example, a delay of three years in completing experimentation, or a reduction of 10 percent in the rate of adoption by farmers, will probably reduce cumulative net returns (benefit minus costs) in present-day values by considerably more than would a substantial increase in the cost of the research itself.

Benefit/cost analysis should include, at a minimum, the sequential estimation of eight distinct characteristics of a research program and its impact (Bottomley 1988). They are: the annual cost of research, its duration, its initially anticipated probability of success, on-farm implementation costs, on-farm benefits, the rate of adoption, the adoption ceiling, and the life of the innovation.

A major difficulty in conducting benefit/cost analysis is in knowing how to define and measure both input and output. A World Bank publication (Schuh and Tollini 1979) has examined this subject in some depth. It stresses that the most important output of the research process is new knowledge, but this is not a quantifiable product nor is there a well-defined market for it, despite the fact that it has obvious economic value. Thus it is not possible to be precise in measuring the value of knowledge to society. Various proxies for output have been examined, including scientific publications, although publication vary widely in the quality and quantity of their content and are not a consistent measurement of value. An alternative approach is to define output in terms of some well-specified innovation, but this suffers from the problem that the level of adoption of an innovation may not necessarily relate to the quality of the research which produced the innovation.

These difficulties in measurement have led to many attempts at evaluation being related to measuring output in terms of its impact on the production process. This enables the role of the program being evaluated in the overall research and development plan of a country to be examined and also permits a look at the adoption rate and its impact on production. Such an evaluation of the impact of research on development is more complex than a straightforward evaluation of implementation. Furthermore it needs to be recognized that the contribution of research to development is influenced by a number of extrinsic factors such as the availability of inputs, relative prices, marketing channels, and extension services. This means that impact evaluation may need to take place some years after the completion of a research activity, and in this sense it is of limited short-term benefit to the research program itself.

Impact evaluations are often used to try to convince policymakers that resource allocations to research represent good investments. Progress in terms of production, income, or marketable produce is usually the yardstick of success used by policymakers, so a high rate of return to research investment can be a strong selling point for research leaders when presenting the research budget for approval by government authorities.

Impact evaluation of research needs to distinguish between the contribution of research and that of other development activities, such as extension, the provision of inputs, adequate infrastructure, and favorable marketing policies. There is a particularly close linkage between education, extension, and research.

Most evaluation studies have been unable to effectively distinguish between these three variables and have implicitly assumed that the returns from agricultural research also include the benefits from education and extension. The few studies that have sought to separate these activities have had to do so subjectively. Because of the lack of sufficient theoretical instruments, these studies have generally attributed most of the benefits to research alone, and in some cases this may have resulted in erroneously high rates of return being attributed to investments in research.

Yet another problem with the measurement of outputs relates to assessing the significance of the maintenance research which is required to overcome obsolescence, particularly that due to changes in disease and pest biotypes. In order to keep pace with these changes, the output of maintenance research may be just as important as that of research on innovation, but identifying and accounting for maintenance research is not an easy task since, for example, the pathogenicity of new strains of pests and diseases does not develop following a predictable pattern.

Problems of measurement can also apply to inputs, although many of these can be defined. However, inputs of highly skilled manpower may be hard to quantify. The pricing and measurement of previous research also represents an input problem unless prior research endeavors are treated as free goods. But this ignores the fact that someone paid for the prior research, and if this cost is not taken into account it can again give a result which places an unduly high social rate of return on the research.

In spite of these problems a number of efforts have been made to use input and output data to carry out benefit/cost analysis of agricultural research. This has been done either in an ex post sense or ex ante.

Schuh and Tollini (1979) have described a number of approaches and models that have been used to make ex post evaluation. They found that there was a rather rich set of research procedures that have been developed whereby research can be evaluated and its contributions and various effects analysed. Different approaches are useful for answering different questions and the particular question posed will vary a great deal depending on the individual problem situation. A major constraint to the use of ex post analysis is, however, the time period between undertaking research and being able to assess its benefits.

For ex ante decision making there is a vast literature from industrial and military research but rather less from agricultural research. A number of models are available with methodological sophistication ranging from simpler scoring models to more complex mathematical programming models. Schuh and Tollini state that the advantages of these models are that they provide a basis for decision making with an eye to the future rather than the past, they pool information from a large number of qualified experts and they provide a means of explicitly relating the research effort to a set of goals. The disadvantages are that those methods which draw on the opinion of a large number of specialists can be quite costly and time consuming and the pooling of a large number of opinions may do little more than to pool ignorance. It is probably for these reasons that the more complicated methods have rarely been used more than once, although selected models may provide a means of feeding some rigorous analytical research into the decision-making process.

Yet another problem in using a formal cost/benefit analysis in a developing country research institution is that the analysis may tend to under-emphasize issues of equity. These may be of considerable importance in countries where income, wealth and power distributions are highly skewed. In many countries food security is the central goal for agricultural research. Busch (1985) has drawn attention to the fact that this has implications in terms of a number of distributive issues such as:

- a) New labor-saving technology may add to unemployment and result in a complete loss of income for certain people.
- b) Labor issues can also be important in terms of the introduction of new crops or varieties that radically alter seasonal labor needs or labor needs that interfere with other essential family activities. Such effects can reduce the demand for casual labor, thereby eliminating the traditional method of redistributing wealth and thereby contributing to reduced food security.
- c) New varieties of crops may encourage the use of marginal lands and lead to environmental degradation, undermining the food security of future generations.
- d) The role of women is significant in that they play a major role in each of the four aspects of household food use: procurement, handling, distribution and consumption. In many countries these four tasks are fully integrated and are part of the daily work activities of women. Changes in the labor patterns resulting from changes in production patterns may disturb this integration and effective agricultural research cannot ignore this intrinsic linkage, which is fundamental to food security.

- e) Broad issues of agricultural research policy may also affect food security, such as the introduction of feed grains or the encouragement of cereals as opposed to grain legumes. These can result in shifts in the protein/calorie balance, particularly of the poorer segment of the population, thereby affecting food security.

#### 1.4 The Clients for Evaluation

The wide range of issues that can be looked at in an evaluation indicates that there can be a broad spectrum of clients for this activity. Since different clients may have different requirements from an evaluation it is important that the specific client be identified when an evaluation process is being structured. The terms of reference for an evaluation review should clearly relate to the needs of the client for that review, recognizing that many reviews will serve more than one client. Among the clients for research evaluations are:

National policymakers, who are interested in the role that research does or can play in national agricultural development. Their interest is primarily likely to be in terms of strategic reviews which define overall research priorities and the resources to be allocated to research within the context of development strategy at large. Such clients are likely to be interested in the potential impact of agricultural research on production and productivity. They may also be particularly interested in the research potential when there is a major change in research capacity, technology potential, world economy or some other factor which modifies the agricultural sector and, therefore, research needs.

Donor agencies, who now play a major role in supporting agricultural research in developing countries. Evaluation reviews provide them with information on the efficiency of a research system and may help them to justify their investments in it or to identify areas needing strengthening in which their assistance programs have a potential role to play. This role may involve either technology and/or management.

Senior research managers, who are able to use the evaluation of past activities to assess the results achieved and to build the lessons of experience into the corporate memory of the institution and thereby to improve its future activities. Such managers have the task of selecting and designing new programs. Evaluation reviews can be used to determine which programs need to be strengthened, modified or deleted. Reviews carried out on a quinquennial basis are of particular value to this group.

Project or program managers within the research system, who can use evaluations of ongoing activities to assess progress, and to identify bottlenecks and problems, so that they can be remedied before causing further damage.

Individual scientists, who are able to use review findings to look at their own research activities within the context of an entire program or institution, and also to place this work into the perspective of adoption and impact as conceptualized by a review team.

Obviously there is a great deal of scope for weighting the scope of a review to cater to one or more of these client groups and the composition of a review team, the nature of its work and the focus of its activities will all need to take this into account.

### 1.5 The Method of Choice

There is little doubt that where appropriate data, skills and time are available, cost/benefit analysis is the method of choice for evaluations whose prime objective is to measure impact. Unfortunately many NARS lack the appropriate data, skills and time to conduct cost/benefit analysis and are likely to be in this position for some time. While external assistance can help to overcome this problem it is unlikely to be able to do so on the sort of sustained basis that effective evaluation requires. Furthermore, even in developed countries there is little practical experience of cost/benefit analysis of agricultural research other than on an 'ad hoc' basis.

Another factor that needs to be taken into consideration when determining how to conduct an evaluation is the fact that, in apparent contradiction to the high returns to research reported from impact studies (Evenson 1987) many program reviews have noted the existence of serious managerial and operational problems with the NARS. In many instances, poor management, rather than a lack of funding, appears to be the principal constraint to research impact. This suggests that in general terms, NARS are better at handling the technical aspects of development than the institutional aspects (Horton 1988).

This view is supported by a study on World Bank projects (World Bank 1983) which indicated that "while the Bank has successfully supported the development of physical research facilities, their success has not yet been matched by improvement in the management of these facilities or the development of institutional arrangements conducive to their proper utilization". In such circumstances cost/benefit analysis is not necessarily the optimal or most cost-effective approach to evaluation in terms of current needs. Simpler evaluation methods such as checklists, would appear, despite their subjectivity, to be more practical and advantageous for many NARS at the present time, (and are indeed the method still used by the CGIAR research centers).

This paper, therefore, restricts its discussion on evaluation methodology to the checklist approach. At the same time it endeavors to recognize the greater conceptual rigor of the benefit/cost approach by placing particular emphasis on the selection of checklist questions which relate to issues affecting resource optimization, although it does not call for quantification in the answers to such questions.

## Chapter 2

# WHAT TO EVALUATE

THE INFORMATION ABOUT A NARS CONVEYED TO RESEARCH MANAGERS AND POLICYMAKERS AS A RESULT OF AN EVALUATION CAN COVER A BROAD SPECTRUM OF ACTIVITIES AND CAN MEET A WIDE RANGE OF INFORMATION NEEDS. THESE CAN INCLUDE INFORMATION ON STRUCTURE, ORGANIZATION, MANAGEMENT, PLANNING, RESOURCE AVAILABILITY AND UTILIZATION, SCIENTIFIC METHODOLOGY AND OUTPUT, LINKAGES TO CLIENTS AND ECONOMIC IMPACT. WITH SUCH A BROAD SCOPE OF WORK POSSIBLE, THE BENEFITS OF EVALUATION ARE LIKELY TO BE OPTIMIZED IF THE CONTEXT OF AN EVALUATION IS DEFINED FROM THE OUTSET. THE USE OF CHECKLISTS CAN BE A USEFUL ADJUNCT WHEN CARRYING OUT THIS TASK. SUCH LISTS CAN SERVE AS THE INITIAL INTERFACE BETWEEN THOSE REQUESTING THE EVALUATION AND THOSE ACTUALLY CARRYING IT OUT.

### 2.1 Introduction

The evaluation of a research institute or program is generally seen as an activity that is not limited to just examining the experimental activities, or their potential or actual impact. A somewhat broader perspective of evaluation is usually adopted, because it has been found that in order to place an evaluation into an appropriate context it is necessary to relate its findings to the agricultural sector as a whole and to the changes taking place in it. It is also necessary to link the research being reviewed to other ongoing and past research activities. Additionally it is usually necessary to take into account the social, cultural, and economic situations in which the research results will be applied, specifically including the policy environment which might influence program implementation or the adoption of research results.

Research cannot be evaluated appropriately without knowing the constraints under which it operates (budget, available manpower, rules and regulations, etc.) and the agricultural context which the research activities are expected to serve. Thus the assessment of a research program needs to include an analysis of the performance in programming and managing the research activities and in optimizing the use of the human and financial resources. In addition it is necessary to evaluate the scientific validity of the research itself and the appropriateness of the research in relation to national needs and priorities.

Evaluating an individual institute, a specific program, or an entire research system, all entail looking at a similar set of characteristics. The difference is one of emphasis. When looking at a research system as a whole, more attention has to be paid to national policy and to economic factors than when evaluating the implementations of a specific program, for which only those policies or economic factors directly influencing the program may be relevant. Similarly, an evaluation for which the prime client is a national policy unit which recommends the allocation of funds for research will need different types of questions asked to one for which the principal client is a research manager interested in manpower utilization or a donor concerned about the use of its technical assistance funds focused on improving the quality of planning and programming.

ISNAR has made several efforts to characterize the various factors which may need to be examined in the evaluation of agricultural research. This has been done by integrating them into a series of subject-oriented checklists for use by an evaluation team. The questions in these checklists serve as guidelines for structuring the content of an evaluation review report (see Chapters 3 and 4) for all types of client.

It should be stressed that at no time ever would an evaluation be expected to address all of the questions in the checklists presented in this chapter. Rather, the organizers of an evaluation would be expected to review the lists during the planning stages and then to select, add, delete, and modify questions with respect to the terms of reference of the particular review under consideration. The checklists are designed to act as catalysts. When deciding whether to include an item on a list, two points should be discussed:

- How would that information be used? It should not be included unless it can be justified by being used. "Being nice to know" or "interesting" is not a valid justification.
- How much work is involved in obtaining it? Whether it is worth the time and trouble of collection depends on how necessary the information is for the final analysis.

In preparing checklists for a specific evaluation review, particular stress needs to be placed on tailoring the content of the review to the practical realities of the program, project, or institute to be reviewed. Human, physical, and financial constraints need to be recognized realistically. Key scientists do move to administrative posts to obtain promotion, funds do not always arrive on time, research stations may be overstaffed to comply with government job creation policies, operational funds may be limited as may be funds for spare parts; budgets may be cut, sometimes irrationally. But these are facts of life in many NARS and any evaluation of performance needs to recognize these realities. This does not necessarily mean that an evaluation should not identify such constraints, point out their significance and try to suggest remedies. But change will not take place overnight and an effective evaluation will recognize this and try to suggest how change can take place within the existing constraints as well as when they are removed.

For purposes of convenience, the checklists which are presented and briefly discussed in the rest of this chapter have been sequentially presented in a normal manner which addresses four major issues with which evaluations have to deal, namely:

- a) The planning, programming, and priority-setting processes - what were the programs' goals and how were they determined?
- b) The financial, human, and physical resources available for conducting the research - how adequate were the resources in terms of availability and utilization?
- c) The results attained and their impact on development - what was the output, was it adopted by farmers and how did it benefit both them and the nation as a whole?
- d) The institutional background - did institutional factors act as a constraint to goal achievement?

## 2.2 Planning, Priority Setting and Programming

The checklist on these topics is designed to bring out what an institute or program is attempting to do and why. The questions start by asking about the agricultural sector as a whole, its status, potential plans, and priorities. The degree of depth entered into in covering such background material will depend in part on the purpose of the evaluation, particularly whether it will be used in communications with:

- 1) foreign colleagues or donors: they may not be fully familiar with background data on the country's agricultural sector and administrative organization and may require such information;
- 2) national policymakers: they may need precise information on overall research budgets and personnel but are unlikely to require national macro-data.

A document prepared for general release may need more background information than will an internal document. The checklist given below is quite restrictive and does not cover key economic and statistical data at the macro-level, nor does it deal with issues such as prices, marketing and credit.

Having asked (however briefly) about the sector as a whole, the questions in the checklist then pass on to the subsector dealing with research. The way in which research priorities are set and resources are allocated need to be established in order to assess their rationale.

The next stage is to move from priority setting at the national level to actual program formulation, how it is done and what programs emerge from this process. The findings from this then need to be related back to the national goals. In some cases a research institute may have a written mandate, or a national program may have a specified goal; in such a circumstance it is important to assess whether the program or institute under review conforms with the specified goal or mandate. These questions can be asked at two levels, first the institute or program as a whole and, second at the micro-level of individual experiments.

Another issue that may need to be addressed at this stage relates to the extent to which the ongoing program structure is influenced (or even distorted) by extra-budgetary support from donor agencies. Their inputs can affect the overall balance of a program in terms of meeting national priorities.

The final questions in this checklist relate the ongoing program to a set of criteria which will permit the economic evaluation of the program. These questions will usually be difficult to answer at this point in time in many NARS, but the mere asking of them should help to start a more analytical dialogue as to why a particular piece of research is being carried out. If, for example, the probability of success is deemed to be low or the lifetime of an innovation is expected to be short, then the opportunity is presented to challenge the rationale for the research in question. Making (and defending) such challenges should be an integral part of the peer review process in any evaluation and can serve as an important training exercise for research scientists and managers.

A SHORT CHECKLIST OF RESEARCH PLANNING,  
PRIORITY SETTING AND PROGRAMMING

1. What is the country's natural resource base?
2. What is considered to be its long-term agricultural potential?
3. What are the principal constraints to achieving this potential and what research is needed to overcome them (in the realm of the program or institute being evaluated)?
4. What are the priorities of the latest national plans for agriculture and for agricultural research and are the latter determined on the basis of available resources, economic analysis, or how?
5. What is the mechanism for deciding the total budget for agricultural research and for allocating the available manpower and funds between the various institutes or programs involved?
6. What are the specific objectives of the program (institute) being evaluated? How were they determined? How do they relate to national priorities? Are they realistic in terms of available resources? Are the needs and priorities of all relevant parties (government policy and planning office, intended beneficiaries of the institute or program, etc.) well articulated?
7. What process is used for formulating the actual work program at the experiment level? Who makes the decisions and does the process involve users? Is the detailed work program all relevant to the institute (program) goals and appropriately balanced in its disciplinary components (including economics)?
8. Has the availability of donor support affected the design or implementation of the national program?
9. What benefits is the research expected to provide to farmers? Can the expected benefits be related to the annual costs of the research, and the likely duration of the necessary research program and also to:
  - a) the probability of success,
  - b) the added costs to farmers of utilizing the research findings,
  - c) the extra farm income which the research is likely to generate,
  - d) the anticipated rate of adoption,
  - e) the ceiling on adoption in terms of land area or stock numbers,
  - f) the lifespan of the innovation?

This list is by no means exhaustive. It offers some theme questions around which others can be developed according to the particular circumstances of the review. Two examples of how this can be done in practice for this and the later checklists are shown in Annex D.

### 2.3 The Resources Available for Research

Having defined and justified the goals of the research it is then necessary to examine the resources available for carrying it out. The achievement of the goals will depend on the availability of appropriate resources and the way in which they are utilized. Most research institutes will claim that they are under-funded, but this is a chronic and widespread problem and, unless an evaluation is specifically charged with making recommendations about funding levels, little may be gained by arguing for greater financial support, for more facilities or more personnel. Rather than doing this it is preferable in a research evaluation to relate the resources available to the programs and goals and to indicate what can be expected from the use of the existing or programmed resources, rather than from hypothetical ones. In this context comparisons between strongly and weakly supported programs or institutes within the same country can be important.

The discussion on resource utilization which follows groups resources under the heads of financial, physical and personnel.

#### Financial resources

Clearly the prime task in assessing the financial resources is to examine whether the budget for the program (institute) is realistic in terms of the program and goals. This topic may need to be examined from two stand-points, the current year and the medium term, since many experiments last for several years. In looking at the budget it is also necessary to examine both the source and the use of funds. The source is important because the funds that come from extra-core (project) monies may be of finite duration, whereas core funding tends to be fairly stable from year to year. The source of funding also relates to its use, since a large part (often 70 - 80%) of core funding may be locked into the payment of salaries. When this sum and the fixed costs for items such as utilities are taken into account, the variable funds available for items such as operational research costs, travel and facilities maintenance may be insufficient to run an effective program. Hence it is not only the total size of the budget that is important but also the relative size of its various components.

In addition to examining the quantitative aspects of the financial resources, an evaluation team may also find it necessary to examine and comment on some qualitative issues relating to financial management insofar as these may act as a constraint to the implementation of the research program. It is not enough to have a budget; it is also necessary to have a system that ensures that the funds voted for research are available when they are needed. While a scientifically based evaluation may not wish to dwell in too much depth on issues of financial management, a minimum checklist should include some questions relating to this subject.

#### A SHORT CHECKLIST OF FINANCIAL RESOURCES

1. What is the budget of the program (institute)? Is it sufficient for the planned program of work? Is it linked to the program through any form of program budgeting system? Is the total budget from all sources consolidated in any formal way?
2. To what extent is there multi-year budgeting? Which aspects of the program are included in this? What continuity of funding exists other than for staff salaries?

3. What are the sources of non-core funding? How large are they? How much do they vary from year to year? What influence do they have on the structure of the research program?
4. What is the budget breakdown between personnel costs, travel, maintenance of facilities, etc.?
5. How large is the capital budget and what is its source?
6. Is the use of the budget constrained by factors such as the timing of its release, tendering procedures, or other factors?
7. How are unexpected cuts (or additions) to the budget dealt with?
8. To what extent do scientists in charge of programs have control over their budgets? Is there a viable mechanism for regularly informing them of the status of their budgets? Or is all control centralized; if so whose hands is it in?

#### Human resources

Manpower is the heart of any research system. An evaluation will need to look first at the human resources available in terms of scientists, scientific and field support staff, and administrators, since without appropriate support, scientists cannot function properly. For the scientists, the disciplinary training and the level of training are also important. Few first degrees give appropriate training for research leadership; on the other hand, much research is routine and every researcher does not need a Ph.D.

Having established what is the human resource base, the next question is to assess whether it is appropriate for carrying out the planned research program. This leads on to questions about manpower training and the adequacy of the existing plans for this. In this context some note may need to be taken of the staff turnover history and the expected future size of the research program staff.

This raises the broader question of whether there is any form of manpower plan and, if so, how it was established, whether it is drawn up centrally or on a station-by-station basis and whether it covers only scientific staff or also technical support staff and managers. In discussing this it may be also necessary to examine the adequacy of local facilities for higher training in some areas.

A review of training needs should not only cover higher-level training. There is also a need in any dynamic research organization for an in-service training program in order to avoid both technical and intellectual stagnation. The extent to which in-service training is practiced should also be critically reviewed.

#### A SHORT CHECKLIST OF HUMAN RESOURCES

1. What are the current numbers of research scientists, field and laboratory technicians, and administrative staff in the program? By levels of training, areas of specialization, and nationality (national and expatriate)?

2. Are the total resources of manpower adequate for a realistic program of research to be carried out on a continuing basis for fulfilling the program's objectives?
3. What are the current numbers of research scientists, field and laboratory technicians, and administrative workers, being trained (in universities) for positions in agricultural research? Is the supply adequate for current and/or projected needs? What is the staff turnover history?
4. Is there a long-term manpower development plan? If so, is this drawn up centrally or by aggregating the felt needs of individual research stations? Does it cover only research staff or institute support staff, including those engaged in management? How was the plan formulated and how does it relate to both local training facilities and the expected availability of funds for training?
5. Is there an in-service training program? How broad is it in its coverage? How effective is it?

There are a number of questions that can also be asked about personnel management, but these would not necessarily be part of an evaluation review. However, where factors such as staff remuneration (both absolute and relative), promotion prospects, conditions of service and so forth have a major influence on staff availability and manpower planning, they may need to be addressed since, unless there is appropriate manpower there is no research program.

#### Physical resources

Information on the number, size and condition of facilities, equipment and supplies needs to be related to that on manpower and training, because equipment is of little use in the hands of staff who are not trained to use it. Likewise trained staff cannot achieve very much unless they have appropriate equipment and facilities to work with.

The checklist here also contains questions on library and documentation facilities and on the location of experimental stations and on-farm research. It is intended to provide information on both the status and the use of facilities, equipment and supplies.

#### A SHORT CHECKLIST ON FACILITIES, EQUIPMENT, AND SUPPLIES

1. Where are the physical plant, experiment stations and on-farm sites located? Are they appropriate in terms of size, number and proximity to production areas?
2. Are the resources of office, laboratory, library, equipment and supplies appropriate for the planned program of work?
3. Is the available laboratory and farm equipment maintained well, are there adequate spare parts, are there staff trained in its use and maintenance? Is lack of equipment a constraint?
4. How well is the library equipped with up-to-date materials in the relevant fields? How accessible is its material to scientists? How much do they use it?

## 2.4 Research Activities, Achievements and Impact

In the pages above we have discussed the evaluation of how the goals for the research were set and the nature and use of the resources available for meeting these goals. We now come to discuss the evaluation of the results achieved. To do this it is necessary to ascertain whether goals were attained and, if so, how this was done and with what implications in terms of both technology and development. Most agricultural research workers are biological scientists for whom this part of the review process, which implicitly evaluates the scientists themselves as well as the quality of their work, is the core of evaluation.

Such an attitude is, however, a simplification since, as discussed in Chapter 1, the technical quality of the research being carried out in most NARS is seldom as much of a problem as the institutional and organizational constraints. This situation is probably the result of the massive efforts made by many NARS during the last two decades in the area of post-graduate training. Little of this training has been devoted to research management and many NARS still perceive the evaluation process primarily in terms of its technical components; thus the economic, institutional and managerial aspects of the task are often downplayed.

Because of this strong technological bias, as well as the high quality of the technology that most NARS are now capable of producing, it is essential that the team evaluating the technology should be competent scientists and have the professional skills necessary to review research quality and methodology. But quality control should only be one component of the review of technology, since it is also necessary to examine the output from the research and the extent to which it is utilized by farmers. There is an important distinction between these two elements. For example, a new technique or variety can represent a positive output, but if it is not adopted by farmers the developmental impact of the output may be zero.

Thus, in order to evaluate the performance of a research unit it is necessary to look at its research activities, achievements and impact. These have been divided into three separate checklists below.

### Activities

It will seldom be practical for evaluators to review the work plans and experimental designs for every experiment in a program or institute, although this should be done for each major program area. At the level of the individual experiment it can be very helpful to trace the life history of a few individual research protocols, preferably selected at random. This can highlight strengths and weaknesses at every level in the implementation of the program, from the criteria taken into consideration by the researcher when designing the protocol, to how the work was actually conducted, the results interpreted and reported.

### A SHORT CHECKLIST ON RESEARCH ACTIVITIES

1. What is the rationale behind the experiment protocol? How does it relate to program objectives? Is it supported by an appropriate literature review?

2. Who planned the experiment? Who was consulted in the process? Is the design appropriate?
3. Who implemented the experiment?
4. Who recorded the results? How?
5. Who reported the results? How? To whom?
6. Who interpreted the results? What statistical techniques were used?
7. Who disseminated the results? How?

Following such a detailed examination of selected individual experiments, the evaluators will wish to piece together a synthesis of the research actually under way and to examine whether:

8. the ongoing program has an appropriate focus and relevance in terms of farmer needs;
9. the activities are being conducted and supervised as planned;
10. problems encountered call for major revisions in the program;
11. the field and analytical methodologies are appropriate;
12. the end results are appropriately reviewed and disseminated; and
13. the results can be expected to have a tangible effect on production.

### Achievements

After conducting such an examination of the research itself, the evaluators are in a position to try to assess what have been its principal achievements. This needs to be done against some form of time-frame. For this purpose a five-year period is often appropriate and is relevant to the suggestion made elsewhere in this paper that reviews of the type being discussed in this chapter can usefully be repeated on a quinquennial basis.

### A SHORT CHECKLIST ON RESEARCH ACHIEVEMENTS

1. What are the principal technological innovations resulting from the research, in terms of products, processes and knowledge?
2. What was the duration and cost of the research necessary to produce these findings, how did this compare with the planned time-frame and cost?
3. To what extent have the accomplishments of the research program been recognized by the government, the agricultural community, agro-industry and the general public? How was this brought about?

In evaluating achievements it is important to recognize the age and maturity of a research program or institute. It is not realistic to

expect many innovations to have emerged from a program in which most of the senior personnel with higher training have been present for a limited number of years. Much will depend on the type of research being carried out. Findings from adaptive research using imported, improved germplasm are obviously quicker to obtain than are results from more basic and long-term activities such as selection for drought tolerance or pest resistance. Evaluators may wish to comment on the achievements of a program in terms of the balance between its long- and short-term research activities, since this can be relevant in terms of creating institutional credibility within a country.

### Impact

Well-planned and executed research leading to new technology or new knowledge may be an acceptable end point for research conducted in an academic institution. But for agricultural research funded by a producer levy or by the public sector, more is usually expected, namely that the research is adopted and has a positive impact at the farmer level. Unfortunately, although the research services are usually held accountable for impact, they have only a limited degree of control on the uptake of their innovations. Changes in farming practices result from interactions between many agro-ecological, technical, social, economic, institutional, and policy variables, of which research is only one.

Thus an evaluation of impact may need to look at factors such as extension, pricing, marketing, labor utilization, and nutrition, all of which can play an incentive or disincentive role to adoption. This all adds to the complexity of an evaluation and may call for more time than is feasible in an evaluation focussed on research, even though the results of such a comprehensive evaluation can be much more far-reaching in developmental terms. Thus, in the checklist suggested below, the depth to which the question (no. 4) on constraints to adoption can be dealt with in a research program review will depend very much on the overall objectives of the review and the resources made available for answering this question.

### A CHECKLIST ON RESEARCH IMPACT

1. At what rate have farmers adopted the new innovations put out by the research program (institute)? Which category of farmers have done so?
2. What is the anticipated ceiling on adoption, and when is it expected to be realized?
3. Are some parts of the innovation packages being adopted at a higher rate than others? Is there any explanation for this?
4. What factors, if any, have constrained adoption? What is needed to overcome these constraints? Are there social or political implications? (These could involve a very broad range of possible factors; agro-ecological, social, economic, government policies, infrastructure, etc.)
5. Who are the principal beneficiaries of the results? How do they benefit? Can this be quantified at either the individual or the national level?

6. How do the end results compare with the ones expected in the causative line from output to purpose and to goals in the program design? (This question will be mandatory if a logical framework analysis is being used.)

## 2.5 The Institutional Background

Apart from considering the objectives of the unit being evaluated, the resources available for meeting these objectives and the results attained, the evaluation process may need to place these findings into an institutional setting, since the institutional background can influence the answers to the questions in a number of the checklists given earlier in this chapter.

Two distinct sets of questions need to be addressed in this area:

- (a) the contextual structure and organization - how the program fits into the national agricultural research system as a whole, and (how) it interacts with other components of the system; and
- (b) the internal structure and organization - how the program is organized, how it is set up internally to carry out its functions and what monitoring and evaluation systems exist to measure progress and performance.

These subjects may be covered quite briefly for an evaluation of a program restricted to one institute but will require detailed attention in the case of a joint activity involving several institutes or in the evaluation of an institute with multiple research activities.

Agricultural development and increased productivity depend on a country's ability to form and manage effective two-way relationships between technology developers and technology users. The success of a research program also depends on the effective operation of linkages with other research organizations, both within and outside the country, with development agencies, and with policymakers.

Thus an important set of questions relating to the institutional background arises from the linkages that the program (institute) has, both with the outside and internally.

The linkages can be official, with formal mechanisms to ensure exchange of information, such as regular meetings, line of reporting, or official visits. These are fairly easy to describe. However, the existence of formal communication mechanisms on paper does not mean that any communication actually takes place and informal mechanisms - communications between individuals rather than between positions - are sometimes the most effective form of communication. Thus this role needs to be considered when evaluating linkages.

### A SHORT CHECKLIST ON THE INSTITUTIONAL BACKGROUND

1. Where is the program (institute) in the national agricultural research system? What degree of autonomy does it have? Does it overlap in its responsibilities or activities with other research units? Are there parallel, interdependent, competing or redundant responsibilities and activities?

2. What linkages exist with policymakers, other related institutes or programs, universities, private-sector research, international agencies and agro-industry? Are such links formal or informal, do they relate to policy, management or technology?
3. What is the level of contact between the researchers and both extension workers and farmers? Do the researchers have an on-farm research program or collaborate in one conducted by extension services? How effective is the information flow from farmer through extension worker to researcher and vice versa?
4. What is the internal structure of the program (institute)? At what level is the research programmed and designed?
5. What is the extent of the dialogue between different research teams and disciplines in the program (institute)? Is there a good seminar program?
6. To what extent is the performance of research activities and personnel monitored and evaluated on a regular basis? How is this done?

In concluding this discussion on checklists it may be useful to reiterate that the lists given here are representative of a spectrum of activities that can be covered in an evaluation. No single review is likely to wish to address all of the questions listed. They serve only as a guideline for use in the review planning process in order to develop a specific list for use in the operational stages of a specific review.

### Chapter 3

## HOW TO EVALUATE

WHEN PLANNING AN EVALUATION IT IS IMPORTANT TO DEFINE WHO WILL USE THE RESULTS AND FOR WHAT PURPOSE. CHECKLISTS CAN BE A USEFUL TOOL IN THIS TASK. FROM THIS STARTING POINT IT IS POSSIBLE TO DRAW UP DETAILED TERMS OF REFERENCE AND TO MAKE DECISIONS ON THE STAFFING OF THE EVALUATION TEAM AND THE TIMING OF THE REVIEW. THE IDEAL-SIZED TEAM IS 4 TO 7 PERSONS. AT LEAST PART OF THIS TEAM MAY CONSIST OF EXPATRIATES, ALTHOUGH THIS HAS MAJOR COST IMPLICATIONS AND IT IS IMPORTANT TO ENSURE THAT THE OBJECTIVES ARE SUFFICIENTLY CLEAR AND IMPORTANT TO JUSTIFY SUCH COSTS. IT IS ESSENTIAL TO STRUCTURE THE WORK PROGRAM SO THAT THE USE OF EVALUATORS' TIME IS OPTIMIZED. THIS IS BEST DONE BY ASSIGNING TO A DESIGNATED PERSON THE ROLE OF TEAM COORDINATOR. HIS/HER WORK MAY NEED TO BEGIN AS LONG AS SIX MONTHS BEFORE AN EVALUATION TEAM GETS INTO THE FIELD. AN EFFECTIVE COORDINATOR WILL ENSURE THAT: APPROPRIATE BACKGROUND MATERIAL IS AVAILABLE ON SCHEDULE, BOTH SCIENTISTS AND EVALUATORS ARE PROPERLY BRIEFED, A COMPREHENSIVE FIELD PROGRAM IS ORGANIZED AND RESEARCHERS ARE GIVEN OPPORTUNITIES TO DIALOGUE FREELY WITH THE EVALUATORS. EFFECTIVE COORDINATION IS A KEY ELEMENT FOR THE SUCCESS OF AN EVALUATION.

### 3.1 Defining the Parameters

When an evaluation is to be carried out, the very first stage is for those with overall responsibility for it to clearly establish who is expected to use the results and what they will use them for. The answers to these questions will determine the appropriate recommendation domains for the evaluation. This will then determine what sort of information will need to be collected and at what level of detail. Thus policymakers will need information on the relative costs and expected returns from proposed research programs, but may have little interest in technical issues. On the other hand, research directors may need detailed technical information if the findings suggest resource shifts.

Likewise, the clients for the research will also determine the level of detail necessary in the data that are collected. There is little point in collecting very detailed data if their use does not form an important component of the evaluation.

The identification of clients and the information they require sets the framework for an evaluation and needs to be clearly specified in the terms of reference drawn up for it. The checklists presented in the various chapters of this paper can serve as a tool in both setting the framework for the evaluation and for drawing up its terms of reference.

### 3.2 Who will do it

Some research organizations, particularly larger ones, are structured to have a special unit for carrying out evaluation. Sometimes this unit is combined with a monitoring unit and sometimes also with a planning unit. However, although planning, monitoring and evaluation have close linkages, and to some degree require common skills, each forms a separate topic on its own. It may be counter-productive to link evaluation too closely to

monitoring and planning, since the evaluation exercise can lose some objectivity if the people carrying it out are seen to be too closely associated with the prior planning and ongoing monitoring of the activity. The optimum approach to staffing an evaluation is to have a mixed team that includes persons from within the program being evaluated as well as outsiders. The former can put the research into an appropriate context locally in terms of cultural, social, economic, and scientific issues. Outsiders may be able to bring independent scientific judgement and, if carefully selected, can also contribute a knowledge of the program area on a global basis rather than purely in the context of the program under review.

If the terms of reference of a review transcend purely scientific grounds and examine issues such as impact, equity and linkages to both extension and farmers, it is important that the personnel of the review team should have some experience in these fields so that the end report can be balanced in its judgements. There is sometimes a tendency to downplay the non-scientific aspects of a review when giving consideration to selecting the evaluation team. Many research directors tend to overstress the scientific reputation of potential reviewers, sometimes disregarding the fact that research impact may be constrained more by institutional, social, economic, or cultural constraints on adoption rather than by the quality of the research. Thus the way in which the research priorities are identified and the research program is formulated may be important issues for a review team to examine.

For this reason the selection of the evaluation team, be they from an internal unit or a combination of national scientists, from within or outside this unit, plus external consultants, needs to be related very carefully to the goals and clients of the review. A good evaluation team will contain an appropriate blend of local know-how, scientific expertise, management skills and practical field experience in both the country of the review and in countries with similar conditions and problems.

In selecting an evaluation team it is desirable to have available the curricula vitae of potential candidates and to equate these with the goals and expected output of the review. Many factors need to be considered when selecting team members: disciplinary and commodity expertise, experience with planning and management, current affiliation, past position in the country, contacts with national officials, and ability to work in an interdisciplinary effort. While it is often useful to have as members of the team some persons who are involved in the program being evaluated, it is essential that some should be "outsiders" (either national or foreign), who can bring fresh questions and views to the evaluation.

As in all team efforts, particular care should be given to selecting the person who will lead the review in order to ensure that he or she has the background and temperament required for this task. The team leader must be sufficiently flexible to be able to work with a team which may be both multinational and multidisciplinary and which may contain people he or she has never previously met.

The actual process for selecting an evaluation team is one to which considerable thought needs to be given. It is important to ensure that the names put forward are not restricted to friends, contacts, or former teachers of the persons working in the program that will be subject to review. Ideally, the names of potential candidates should be sought not

only from the staff of the program to be reviewed, but from other research personnel, both in the government and in academia. Technical-assistance personnel working in the country and visitors from the IARCs also serve as potential sources of names. The research director, who makes the final choice of the review team, should have more than one name for each potential place in the team and should have C.V.s, even brief ones, for all candidates.

If a team is to comprise both national and external personnel there should be no problem in ensuring that there is specialized knowledge of local customs, practices, and constraints on the team. However, wherever possible it is desirable that the non-national members of the team should have some prior first-hand knowledge of the country, or at least of the geographic region, so that they can put their findings into an appropriate perspective in the context of both national needs and the constraints under which the program scientists are working.

### 3.3 How Much will it cost?

The major cost element in an evaluation review is the cost of personnel. For national staff this cost may be quite limited. It will need to cover in-country travel and per diems, plus consultancy costs for persons who are neither members of the organization being reviewed nor public servants. For non-nationals, costs will include international and local air fares, per diems, and consultancy fees. A cost may also be incurred for secretarial assistance and for preparing and printing a report, although this is often absorbed in the operational costs of the unit being evaluated.

Most research evaluation of either programs or institutes lasts for between two and six weeks. The size of the team carrying out an evaluation is usually in the range of 2-10 persons. In the ISNAR experience the ideal team size is between four and seven persons. Where a team is partially national and partially external, there is considerable merit in pairing the reviewers and having a team of two or three nationals and two or three expatriates. In addition, it is very useful to have a national person act as coordinator to the evaluation team and to be responsible for its logistics and administration. In Annex B an indicative costing is shown for an evaluation review carried out by a seven-person team. The estimated cost is US\$ 44,000 for a team containing 3 overseas consultants and about half this sum for a 5-6 person team comprising only national consultants. While a cost of US\$ 44,000 may seem high in absolute terms, it needs to be placed into perspective in terms of the potential benefits that can arise from the better research planning that can result from a review. A well-structured and conducted review can have a number of spin-off benefits in terms of improved methodology, management, and training. The research manager concerned about the costs of an evaluation, particularly if it is being provided by donor funding, may wish to compare the potential benefits against those likely to result from short-term technical assistance, which a number of donor agencies now cost at US\$ 10,000 or more per person month.

In a sense, an evaluation review is a quality control exercise, and it is a policy decision for the research leaders as to how much they wish to spend on it. While there are no generally accepted norms for this exercise, the review costs detailed above, if applied on a quinquennial basis to a research institute or program with an annual budget of

US\$ 1 million, would represent 0.9% of the annual program budget. Whether or not this is an appropriate level of expenditure on this type of quality control exercise, is for the national research leaders to decide.

### 3.4 When will it be done?

If evaluation is to be effective it needs to be done periodically at different levels. The higher the level the less the frequency; thus reviews that look primarily at the relevance and scope of programs and projects in relation to national priorities and institute mandates are probably best done at a frequency of 5 to 7 years. The breadth of such reviews and the limited time available for studying background material often mean that they cannot go into details of individual research or scientists. Institute or program reviews may permit the latter to be done, or at least some selective samples to be examined in depth, and may need to be done more frequently.

The degree of emphasis on macro as opposed to micro activities may also be related to the degree of maturity in the institute or program being evaluated. In a relatively young program, given the time for research findings and impact to filter through, the most important priority may be to ensure that the objectives are oriented in an optimal manner, whereas for a well-established program the need may be greater for evaluation of the technical quality and management of the ongoing operations or the impact of the results.

Whatever the orientation of the evaluation, the personnel involved will need to recognize that much current agricultural research is multidisciplinary, often location specific, and conducted in the field so that it is subjected to the vagaries of the weather.

Multi-year experimentation is required to even out environmental influences and experiments have to be designed to cater for farmers who are numerous, scattered, unorganized, difficult to reach, and of varied socio-economic backgrounds. Research is also usually a public-sector activity with all the built-in constraints of bureaucratic delays and hurdles. These factors all imply that evaluation can be a difficult (and costly) procedure. It can also be disruptive for the staff of the program under review who may be asked to prepare background material and be subjected to intensive questioning by evaluators.

Evaluations often lead to changes in program structure or resource allocation; such actions can disturb the existing pattern of work or the structure of an institute. Thus evaluations are often not looked on kindly by researchers. They should, therefore, not be carried out more often than necessary. An evaluation every five years is probably a useful approach for performance monitoring in general. Additional evaluations may be called for when an institute or program has obvious problems, when major changes in government policy take place or in other special circumstances.

Apart from the question of frequency, consideration also needs to be given as to the specific time of year when an evaluation should be carried out. The criteria that may need to be considered here are firstly the cropping cycle and secondly the time when the evaluation report is required. With respect to the former it is clearly preferable to conduct the evaluation of an institute or a program involving an annual crop at a time when that

crop is in the ground, so that field experiments can be examined in situ and features such layout, design and husbandry are easy to examine. Likewise an irrigation program is best reviewed when the irrigation system is working and a livestock program when experiments are in progress.

The comment with respect to report timing relates to the fact that if an evaluation report is needed to help restructure a research program it is important that the report should be available at the time when decisions about restructuring or future resource allocations are being made. Since planning and programming usually takes place on an annual or five-year cycle there is little point in presenting a comprehensive evaluation report with many recommendations for change immediately after a long-term development plan has been tabled and agreed. This means that an evaluation report may have to comply with certain deadlines. This may be easier if the evaluation process is institutionalized in some way, either through the existence of a structured evaluation unit in the research organization or by being formalized as part of the programming process.

Institutionalization should not present difficulties in many organizations where long-term development plans are formulated on a recurrent five-year basis and where research program reviews may be linked to them and also be quinquennial. In such circumstances there is considerable merit in having a program evaluation report available about one year before the start of a five-year plan. There is then sufficient time to build the evaluation findings and recommendations into the strategy for the next five-year development plan. Such a procedure can be institutionalized without necessarily institutionalizing the formation of an evaluation unit.

### 3.5 Planning

To conduct an effective evaluation requires a great deal of advance planning. Ideally this will need to start three to six months before the evaluation proper commences. The reason for such a long lead time is that persons with the appropriate experience to serve as both national and international consultants on an evaluation team often have many commitments and may not be available at short notice. Packaging several such persons into a team that can carry out an evaluation at a specified time usually requires a lead time of several months. Postal delays and travel commitments of potential team members can mean that negotiating with them can take several weeks and this time period can be lost if a person is found to be unavailable. Thus it is advisable to make contact with potential evaluation team members about six months in advance of the evaluation date.

Another reason for commencing planning well ahead is that the preparation of background material by the head office of the research agency and by the institute or program being evaluated comes as an additional task to their regular work program. It is likely to be carried out more willingly and comprehensively if it can be done over a period of several months rather than at short notice. A long lead time for the preparation of background material may also make it possible to distribute this material to members of the evaluation team a few weeks in advance of the evaluation, thus giving them more time to read it. Where background material can be made available in this way, appropriate provision can be made in the contractual terms offered to consultants to enable them to carry out some pre-reading as part of their consultancy assignment.

Another long-term issue in planning for an evaluation is the need to draw up the appropriate terms of reference well in advance. This subject is discussed more fully in the following chapter of this paper, but it is relevant to mention here that it is highly desirable that when potential members of an evaluation team are asked to indicate their availability they should be given the full terms of reference for the evaluation. This permits the individuals concerned to assess their own suitability for the task. Individuals asked to evaluate a program in their general area of interest and expertise may not always be interested, or indeed competent, to do so if the evaluation has a particular bias towards evaluating for example, adoption rates, developmental impact or social considerations, as opposed to scientific quality. Thus the full terms of reference should be distributed to potential team members before they are contracted. For preparing such terms of reference the checklist approach offers a useful methodology.

Another feature of the planning task is to brief the head of the program or institute being evaluated and his or her senior staff as to what objectives of the evaluation are, who will be doing it, and what contribution is expected from the institute or program staff. This can be a very important task in building support for the evaluation exercise, as being something that will help to strengthen the institute or program rather than being a surveillance or policing activity. The former approach is only possible if the staff concerned recognize that they have a participatory role and are expected to have a full dialogue with the evaluators. This, of course, requires that the evaluators should be persons whose knowledge and experience qualifies them to carry out a peer review.

### 3.6 Coordination

Given that the preparation, conduct, and follow-up to an evaluation review comprise activities that are likely to last for several months, it is desirable that one individual should be concerned with their coordination. Where an institute or organization has a formal evaluation unit there should be no difficulty in designating a member of this unit to carry out this task. Where a formal unit does not exist, it is necessary for the agency or person with overall responsibility for the evaluation to assign an individual to this task. Ideally the individual should come from the research agency itself and should be of sufficient seniority that he can deal directly with the head of the unit or program being evaluated and also with the personnel carrying out the evaluation.

The first task of such a person will be to bring together the persons responsible for agreeing on the terms of reference and selecting the evaluation team. He or she will then need to coordinate the recruitment process and to organize the preparation of background material for the evaluation team.

In consultation with the respective directors, the coordinator will then need to assure that the program for the evaluation is agreed and that the necessary logistics are organized in terms of provision of transport, accommodation and appointments for the evaluation team. Appropriate secretarial support and logistic backup for preparing the report and printing will also need to be assigned.

The designation and mode of operation of the coordinator will obviously vary from country to country, but in all cases it is important that he or she should enjoy the confidence of the head of the organization sponsoring the review and should have direct access to him/her at all times.

### 3.7 Preparing Background Material

As mentioned previously, there is considerable merit in having appropriate background material prepared in advance of a review so that it can either be circulated to evaluation team members by post or be available for them when they arrive to commence their work.

It is also useful to have the coordinator devote some thought and attention to defining what background material is most suitable with respect to the terms of reference of the evaluation. It is necessary to strike a balance between having very little available as opposed to giving the evaluation team a large mass of published material and unpublished working notes and documents which will require of them many hours of reading. In the context of defining appropriate background material, a checklist approach and the details provided in the terms of reference for the evaluation are of prime importance.

Most research organizations have (often stated in national five-year development plans) precise statements of their mandates and goals. Most also have official programs of work and budget statements which detail their financial and manpower resources and the key facts regarding their laboratories and research stations. It is also customary to have a work program in each institute or department, and very often there is an individual protocol for each research program and each experiment. A review team will need to synthesize this material into a set of concise and informative tables. To this end it is useful for the coordinator or his office to prepare summary tables ahead of the team's arrival. The evaluators may wish to go into more detail on certain topics, but the less time that they spend in hunting down information and preparing tables, the more they will have for talking to scientists and for looking at the research program in the field.

Depending upon the type of evaluation and its objectives, there may also be a need to provide the evaluation team with some macro-data covering the pattern of crop or livestock production in a country and also to make available data on past trends in both demand and supply. This sort of information is particularly important in terms of evaluating whether a particular commodity appears to be receiving adequate, or possibly too much, attention within the framework of the national resources available for agricultural research.

No two reviews will require the same background tables, so that it is not possible to offer a 'master' checklist in this context. For illustrative purposes Annex C lists the tables prepared as an annex to one evaluation report with which ISNAR has been associated, but this should be regarded only as an example, rather than as a model.

In preparing background material for an evaluation team, it is important to try to ensure that such material is as consistent as possible. Related information may be gathered from several sources and may not be consistent. An evaluation team is unlikely to have the time or to be in the position to make decisions as to which are the best set of data. Considerable time

and effort can be saved if they are given a consistent set of data. In order to do this it is important that the person providing the data to them should analyse them impartially and avoid introducing their own biases where conflicting data exist.

### 3.8 Organizing Field Visits

Another task the coordinator needs to carry out ahead of the arrival of the evaluation team is that of organizing their field program. In a small country this may be easy and there may be no difficulty in a team meeting with all the key scientists and visiting most of the research locations. This may not be the case in a larger research institute or program where the field locations to be visited may need to be selected carefully.

The evaluation team should meet with as many researchers as is practical and, depending upon the objectives of the review, it may also be important for them to meet with policymakers, educators and extension staff. It is probably less important to meet with farmers, unless this is done in conjunction with extension staff, since research normally passes to the farmer through the extension services. Where farmers are visited, an effort should be made to visit a selection of farmers with different levels of farming and adoption. An evaluation team should not go only to 'pilot' or 'show' farms.

If a research institute or program covers a wide range of different geographical areas or eco-systems, it is important that the evaluation team should visit a selection of these in order to gain an understanding of the ecological adaptability of new materials being generated by the research.

As far as possible, a written itinerary should be prepared in advance of field visits and all persons to be visited should be advised about the visit proposed to them so that they can organize their presentation. It is usually only possible to visit individual scientists or groups of scientists for a limited amount of time, and when they are advised about the visit it is desirable to offer them some guidelines as to the particular issues that they should bring up during their time with the team. Particular effort should be made to try to avoid each group of scientists presenting to the evaluators detailed background material which is common to many aspects of an institute or program and which the evaluation team will already have received in their written background brief.

### 3.9 Briefing the Team

The first day of an evaluation review should be devoted to presenting the team with background material, if this has not already been sent to them by mail, and to briefing them by the senior staff of both the sponsoring unit and the unit being evaluated.

Before the whole team meets, the person designated as team leader should have a brief meeting with the coordinator and the head of the sponsoring agency to review the objectives of the evaluation, confirm the terms of reference and agree on the broad outline of the field program. The whole of the evaluation team should then meet to be briefed by both the sponsor and the unit being evaluated. This briefing should involve a background presentation describing the organization and role of the institute or

program under review, its relevance in the national research scenario, the rationale for the evaluation and the specific objectives and terms of reference of the review. This presentation should usually be chaired by the (Permanent) Secretary for Agriculture or the Director (General) of Research. It should be followed by a presentation by the unit under review in which its leader and section or departmental heads describe briefly the objectives of their programs, the resources available to them and the results achieved to date. Much of the material may be new to some members of the evaluation team and it is important to recognize that too much detail at this stage of an evaluation can be counter-productive. The purpose of the initial briefing session should be to give a broad general background. It is preferable that such a session does not last more than about four hours. The rest of the first day of the review is best allocated to collecting and reading background material and to organizing the logistics of travel, hotels, payments and so forth.

It is also important that the team should have time for informal interaction among themselves. If the team consists of both national and overseas personnel it is desirable that during this period they should all be together, even if this means that the national staff move into a hotel, since during this formative part of the review there is likely to be considerable informal discussion regarding structuring the work program. It is during this period also that the team leader needs to assess how he will allocate reporting responsibilities. In doing this, he should clearly specify the role of each member of the team. The division of labor during fieldwork should be clarified, as well as what is expected from each member: only notes and conclusions, the draft of a section of the report or the final text.

Even with an experienced team of evaluators, a fairly structured set of meetings of the entire team is desirable before the fieldwork begins. Such meetings should include on their agendas:

- discussing the background information available, key issues identified, and the preparation of the detailed scope of work. This is already an analytical discussion, during which relevant experience from other programs and countries can be mentioned.
- drawing up a list of information still required.
- discussing the division of responsibilities within the team and the logistics for site visits (how the team will divide).
- discussing the (tentative) format for the report (see chapter 4 for an elaboration of this topic).

### 3.10 The Method of Work

The program of work for a review team will, to a large extent, have been laid down when its visits were being organized. It will generally consist of visits to different research locations, both research stations and experiment farms as well as farmers' fields. On these visits, laboratories, greenhouses, barns and fields will be visited and experiments seen and discussed. In addition, research staff will make formal presentations and offer overheads, transparencies and tables to describe their work and results. The meetings which take place may involve anything from two to three up to a hundred or more people.

It is helpful to begin each part of the fieldwork with a formal briefing by the leader and senior researchers present. This serves several purposes. For the researchers, it is an opportunity to behave as partners to the team rather than as an object of evaluation. For the team, observing the interaction among the research staff (who speaks, who sits where) and which issues they choose to emphasize, is as enlightening as the content of the briefing itself. The briefing should not be limited to a description of the program structure and resources but should present current research activities and results to date, as well as eventually the researchers' views on problems and constraints to achieving objectives. What is left unmentioned, even though the team may be aware of it, can be as important as what is presented.

A visit by evaluators is often disruptive, but it can be particularly annoying to researchers and program leaders to be asked to describe a program that is already well documented. Questions about clear and absolute facts should be answered from available documents. It is a waste of time for all concerned to use interview time or field visits for this. However, not all facts are "clear and absolute", and data about budget disbursement, staff involvement, and experiments in progress cannot simply be taken from program documents, since they may differ from what is planned.

In general, more is achieved when only a few people are present, and in many cases it is more productive to have the evaluation team split into groups of one or two in order to meet with individual scientists, once introductory presentations by group, department or section leaders have been completed. In some institutes there is a tendency for department heads to dominate the meetings and to do most of the talking. But department heads seldom know details of all experiments and by suppressing contributions from junior staff members they deprive the evaluators of the opportunity to assess the capability and creativity of staff members at large. Thus, every effort should be made to have the review as participatory as possible. This also permits the review team to get a better picture of constraints, management, training and other aspects of institutional capability.

There is, however, a danger in having the evaluation team work in a way that is too fragmented, since this can result in duplication of some activities and also a failure to adequately cover others. It is, therefore, desirable that team members should not work in isolation for more than 2-4 days at a stretch without coming together with at least some of their colleagues in order to share views and experiences. This is particularly important from the standpoint of quality control, in order to ensure that there is a modicum of consensus among team members in terms of the research operations that they examine from different perspectives, and on which they may be offered different viewpoints by different researchers.

### 3.11 Linking to Other Scientists and Potential Users

In order to place research activities into context, an evaluation team may need to meet with local officials of the extension service, development agencies and other entities involved in research and the adoption of new technology. It is also useful to speak with farmers, although in practice it is often difficult to do this because the research institute may simply select "model" farmers who, in the presence of the researcher/extension staff, may do no more than reiterate the farmer's understanding of official extension policy. The same problem often arises when extension service personnel are visited in a group.

In practice ISNAR has found that it is not easy to combine a review of research activities with an analysis of the way in which the research service links to others. In many countries there is still a dichotomy between research and extension. The program prepared for an evaluation team with respect to examining communication linkages often serves as no more than a starting point for an evaluation in this area. Better results may be obtained by requesting unscheduled stops and visits during the course of field travel in order to try to obtain spontaneous responses from farmers and local extension agents. However, in many countries even this is not satisfactory, since questions and answers will have to be translated for overseas team members by the accompanying officials of the research or extension agency. In many cases all that can be done is to obtain a very general impression of the way in which research is linking to other scientific activities and to its potential users. If this topic is one of paramount importance it may need to be reviewed as an entity on its own.

### 3.12 Assessing the Impact

It is often argued that impact is too complex a criterion to be used for the evaluation of agricultural research, since agricultural development is not a function of technology alone. This fact is often used as an alibi for poor-quality research and for technology which falls by the wayside on its journey from the research institute to the farmers' field. Poor extension technology is often blamed for the low adoption rates of new technology, when in reality it is the low profitability or impracticability of the new technology which is the reason for the low adoption. Thus it is important to evaluate research based on the utility of the results to the end users. But, as discussed in chapter 1, measuring such utility is not easy.

For this reason a great deal of effort has gone into the selection of simple impact indicators and the means of verifying them. This has led to the widespread use of the logical framework, which has been referred to in chapter 2 and is illustrated in Annex A. Where such a framework has been established, and this is not common in NARS, an evaluation team will have some guidelines in the form of indicators and means of verification. In most cases this will not be the situation. Evaluators will have to use highly subjective judgements. Amongst the indicators which are often used to try to measure the impact of research are:

1. The number of research publications. This is not an acceptable measure of impact on farmers, since publications do not necessarily lead to changes in technology use.
2. Adoption rates. These are also not a measure of research impact so much as one of a successful extension service. In any one year a number of farmers will be trying a technology for the first time. Adoption also does not represent impact unless it results in increased productivity as a result of adoption. A better indicator, but harder to measure, is the number of repeat adopters.
3. Production. This has to be interpreted carefully because of year-to-year effects and because it does not take account of costs or stability. Productivity is probably a better index, but harder to measure. Increased yields on experiment stations are not a measure of impact and should not be extrapolated to estimate on-farm production, although this is commonly done.

4. Land use. The time trend in cropping intensity or area of land placed under a given crop or variety can be a useful measure of impact for research results, since it can indicate which opportunities for production have been added to the alternatives available to farmers.

A large number of proxy indicators are used to estimate changes resulting from the introduction of technology. These may involve estimates of household net income, household well-being or quality of life, balance of payments, trade data, job creation and so forth. But all these pose many difficulties in both definition and measurement and many of them may be influenced by factors external to research, ranging from the weather to the availability of inputs, quality of extension and marketing opportunities.

Given such problems, it is questionable whether, in many circumstances, the logical framework has any advantages over the subjective judgements of an experienced and well-balanced peer team of evaluators, particularly if the questions posed to these evaluators are placed within an appropriate framework which will provide valid answers to key questions such as whether:

1. the objectives and goals originally established were clearly defined and actually met;
2. the results of the evaluation can be fed back into the research process to improve research allocation and the establishment of priorities; and
3. the evaluation provides information to help justify the allocation of resources for the research programs being undertaken.

These are the sort of questions that research directors and policymakers are interested in. They usually feature prominently in the terms of reference for research evaluations. Appropriate checklists can serve as a useful tool in providing answers to such questions.

## Chapter 4

### WRITING THE REPORT

THE END PRODUCT OF AN EVALUATION IS USUALLY A WRITTEN REPORT. THE IMPACT OF THE EVALUATION WILL, THEREFORE, DEPEND LARGELY ON THE WAY IN WHICH THE REPORT IS PLANNED, STRUCTURED, WRITTEN AND USED. THE REPORT SHOULD SERVE TO ANSWER THE TERMS OF REFERENCE GIVEN TO THE EVALUATION TEAM AND THESE SHOULD, THEREFORE, BE PRECISE AND EXPLICIT IN TERMS OF ADDRESSING WHAT ARE SEEN AS THE KEY ISSUES TO BE EVALUATED. CHECKLISTS CAN PROVIDE GUIDELINES FOR DRAWING UP THE TERMS OF REFERENCE. THE REPORT ITSELF SHOULD BE KEPT AS SHORT AS PRACTICAL, WITH BACKGROUND TABLES FORMING A SUPPLEMENTARY VOLUME IF NECESSARY. REPORT FINDINGS SHOULD BE PRESENTED VERBALLY BY THE EVALUATION TEAM TO THE CLIENT AGENCY, AND THE UNIT THAT HAS UNDERGONE EVALUATION SHOULD RESPOND ON A POINT-BY-POINT BASIS TO EACH RECOMMENDATION FOR CHANGE. PROGRESS IN THE IMPLEMENTATION OF RECOMMENDATIONS SHOULD THEREAFTER BE MONITORED ANNUALLY AND BE REVIEWED IN DEPTH WHEN THE RELEVANT UNIT IS NEXT EVALUATED.

#### 4.1 Planning the Report

The end product of an evaluation is usually a written report. The value and impact of the evaluation will, therefore, depend considerably on the format and content of this report. The users of the report will be policy-makers and research managers (and possibly donor agencies). Such people are usually overburdened with reading. For this reason it is desirable that considerable attention be given in preparing evaluation reports to ensuring that they are of manageable length and do not repeat a lot of descriptive information that is already known to their intended readers.

The report format should be one that answers, in a concise and constructive manner, the questions posed in the terms of reference. Where a report is intended to have an external circulation (e.g., to donor agencies) some background descriptive material may be needed but this should be kept to a minimum, if necessary by citing background documents rather than quoting them. Research managers and national civil servants rarely need to be given lengthy descriptions of the status quo situation in their own country or organization. What they seek from an evaluation is an analysis of strengths, weaknesses and performance, with appropriate recommendations for change.

Tables and figures should also be restricted in the main report text and, where possible, should be put into an annex or a second volume. Here again the number of pages should be watched and only tables or figures that contribute meaningfully to the text should be included. A report of 100 pages of text and tables will be more widely read than one of 200 pages, and 50 pages will be scrutinised more carefully than 100 pages. This is not to say that there should be a rigid limit on the length of a report; an evaluation of a large institute may need a long report but such reports should be taken as exceptions rather than as the rule.

It is useful to place an executive summary at the front of the report. This executive summary should be targeted for top executive (or

ministerial) reading and should highlight the conclusions and recommendations in as few pages as possible, certainly a maximum of 5 or 6.

#### 4.2 Terms of Reference

The preparation of an evaluation report really begins when the terms of reference (TOR) for its authors are drawn up. The TOR are the instructions to the evaluation team as to what they are expected to do, why they are doing it and who they are doing it for. If these issues are not spelled out clearly in the TOR, the end report may fail to address the right problems or the intended clients. The checklists offered in Chapter 2 can be helpful in identifying issues which may or may not need to be addressed in drawing up terms of reference.

The TOR need to specify whether the evaluation team is looking at the issues of technology quality, management, adoption rates, available resources and so forth, in order to guide the reviewers in their task. Likewise, the terms of reference need to specify the clients of the report (which may have multiple clients), since the level of interest and comprehension in subjects such as technology, finance and management will vary amongst research managers, administrators, politicians, donors, etc. Good terms of reference are, therefore, essential for a good report.

It is common for the TOR to have two components, general and specific. The general terms of reference should be brief and concise and should summarize the broad goals of the review. In many circumstances it will also be desirable to add to them some specific terms of reference which either detail the issues that the evaluation team is required to examine or provide an outline of the report that they are expected to produce.

General terms of reference drawn from some reviews of research programs and institutes in which ISNAR has participated are shown below.

#### GENERAL TERMS OF REFERENCE

1. The evaluation team will review the activities and management of the program of (the research institute XXXX).
2. The primary purposes of the review are: (a) to provide the Government of country AAAA, its research leaders and the management of research institute XXXX with an analysis of the past, ongoing and proposed activities of the research program; (b) to identify ways and means of strengthening the research program; and (c) to increase the in-house evaluation capacity within XXXX.
3. The review will report on the past, existing and proposed programs of the research institute and will make recommendations with respect to:
  - (a) their management
  - (b) the quality and relevance of the current and proposed research
  - (c) the adequacy of the human, physical and financial resources
  - (d) the effectiveness of the linkages of XXXX with the scientific establishment both in AAAA and overseas
  - (e) the nature and effectiveness of the linkages with the extension services and other agencies providing services to agriculture
  - (f) possible new areas at XXXX for national, regional and international support.

### SPECIFIC TERMS OF REFERENCE

The specific terms of reference are intended to advise the evaluation team of particular aspects of an institute or program, such as the link to extension or the use of a particular technique or germplasm source that they should examine during the course of their work. The specific terms of reference may also be used to indicate a reporting outline into which are built a number of questions selected from the checklists in chapter 2. The selection of questions will obviously vary from institute to institute or program to program. The two examples shown in Annex D are taken from actual evaluations with which ISNAR has been associated. In each instance the questions listed were developed from the checklists but adapted to the specific circumstances of the individual reviews.

The questions posed were intended to identify issues which the evaluation team might (or might not) wish to discuss rather than requiring simple yes/no answers. Thus the questions were presented in order to offer suggestions as to the nature of the report content. They indicated that answers were required to a number of programming and management issues as well as technical ones. In this sense, offering a report outline as part of the specific terms of reference provided a guideline to the evaluation team as to the issues that they were expected to address. This helped to concentrate the focus of the report on the issues of prime interest to those who were paying for it.

#### 4.3 Writing the Report

It is preferable that a first draft of the evaluation report should be written and completed while the whole evaluation team is together in the field. This is particularly important when the team includes consultants from overseas who will leave the country after the review is completed. To finish the report while in the field requires meeting rigid deadlines which may be difficult for people not writing in their first language. The reporting process is facilitated by having a clear outline and by designating specific sections of this to individual team members at the start of the review. Discussion of the report is likely to be facilitated if it has a clear system of headings and subheadings and if these are given a consistent system of numbering for easy reference. Recommendations may appear in the text but should be pulled together in an introductory or final summary chapter in which they are numbered sequentially. This allows for easier discussion of the report findings.

The conclusions and recommendations of the report need to be agreed by the whole team meeting together so that the recommendations can be presented verbally to the key personnel in the sponsor agency at a meeting which should also be attended by the entire evaluation team. This allows for the national research organization to advise the team of factual errors and for these to be corrected. It also permits sensitive issues to be discussed in a small group. It has to be recognized that evaluation does involve subjective judgements and where these are critical of current practices there may be sensitivities, so that it is important that any criticisms should be couched in a manner that can be interpreted constructively.

Following the verbal presentation to the sponsor institution of the principal findings and recommendations of the evaluation, the evaluation

team should agree on any revisions that need to be made. The team leader should then edit the report for consistency and style and also to take account of the sponsors' initial comments. Detailed editing can be carried out most cost-effectively by the team leader working alone after the main report is completed. But this requires that each team member hands in a draft which, while not necessarily complete in editorial terms, is finished in substantive terms. This should be done before the team disperses at the end of the evaluation.

The completed draft report should then be sent to the members of the evaluation team and to the sponsoring organization. Both should be given a four-to-six week deadline for submitting comments. Once this deadline is reached, the report should be finalized. Alterations to the draft report should be principally factual, and substantive issues should not be changed. The team should have agreed on its main recommendations and sounded out its sponsors on them when they meet together to discuss the findings of the evaluation. On this schedule a final report should be possible six to ten weeks after the field visit is completed.

#### 4.4 Using the Report

What happens to the final report lies largely in the hands of the agency to which it is submitted. The fate of a report often depends on who originally commissioned the evaluation and on the care with which they drew up the terms of reference. If evaluation is seen as a management tool rather than as a surveillance activity, then it is likely that the sponsoring agency and the relevant research directors will be presented with findings that they can use to strengthen the research institute and its programs.

Indeed, the whole premise of this paper is that evaluation should be seen as a management tool and if the terms of reference and the report outline are drawn up to reflect this the ensuing report should provide management with valuable information. But for this information to serve any real purpose it has to be used constructively. To do this there is a need for the national research managers to systematise how they handle evaluation reports.

One useful way of doing this is to invite the institute or program that has been evaluated to submit a written response to each of the recommendations in the evaluation report. These responses should be numbered to conform with the numbering of the recommendations and should indicate if and when the recommendations will be implemented and if not, why not. Such a commentary should be prepared within a specified time (say two months) of the final evaluation report being tabled and should be sent to the senior government official responsible for research (director general, permanent secretary, etc). Particular attention should be given to commenting on report recommendations that the recipient institution or program does not wish to implement. A meeting chaired by the National Director of Research or his representative may be necessary to define national policy on any disputed recommendations.

To ensure that the response of the institute or program that has been reviewed is more than just a token one, it is also desirable to monitor the progress made in implementing agreed recommendations. This is best done annually. In the long term the recommendation list and institute responses should also serve as essential background material for when the institute or program receives its next evaluation.

Where the evaluation only covers part of a research organization or agency, there are opportunities for using evaluation reports and institute responses as staff management training material. Apart from routinely distributing this information to other institutes or programs in the research agency for information purposes, evaluation reports and responses can provide useful case-study material for in-service management training for middle- and senior-level staff. In order to do this, some form of institutionalized further education program will need to exist to ensure that valid training material is utilized appropriately. If this is done, evaluation can be used as a management tool, not only for current management, but also for the next generation of managers.

BIBLIOGRAPHY

- Bottomley, A. (1988) Evaluating Research Programs, ISNAR, P.O.Box: 93375, The Hague, Netherlands, unpublished mimeo, 145 p.
- Busch, L. (1985) Evaluating Agricultural Research Institutions in Developing Countries, Dept of Sociology, University of Kentucky, U.S.A. mimeo 45p.
- Dagg, M. (1986) Evaluation of Agricultural Research, Paper presented at "Workshop on Improving Agricultural Research Organization and Management", The Hague, 8-12th Sept.1986, ISNAR, (op.cit.) mimeo. 16 p.
- Evenson, R. (1987) The International Agricultural Research Centers, CGIAR Study Paper no 22, CGIAR, World Bank, Washington, D.C., U.S. A., 73 p.
- Horton, D.E. (1988) Assessing the Impact of International Research, Paper presented at the ISNAR/Rutgers "Workshop on Methods for Assessing Research Impact and for Diagnosing Research Constraints", Rutgers University, New Jersey, 6-8th July 1988, mimeo 37p
- McLean, D. (1988a) The logical framework in Research Planning and Evaluation, ISNAR Working Paper no 12, ISNAR, (op.cit.) 11 p
- McLean, D. (1988b) Monitoring and Evaluation in the Management of Agricultural Research, ISNAR Working Paper no 14, ISNAR, (op.cit.) 29 p
- Murphy, J. (1985) Evaluative Activities for Management and Planning in Agricultural Research, ISNAR, (op.cit.) unpublished mimeo 167 p
- Schuh, G.E. and Tollini, H. (1979) Costs and Benefits of Agricultural Research: The state of the Art. World Bank Staff Working Paper no 360, World Bank, (op.cit).
- Shumway, C.R. (1983) Ex-ante Research Evaluation: Can it be Improved. Agric. Admin. 12 91-102
- United Nations (1984) Report of an ACC Task Force on Rural Development (cited by Dagg 1986)
- World Bank (1983) Strengthening Agricultural Research and Extension: The World Bank Experience, World Bank (op.cit.).

ANNEXES

- Annex A An Example of a Logical Framework for a Research Program
- Annex B Indicative Costs for an Evaluation
- Annex C Sample List of Background Tables
- Annex D Examples of Report Outlines

Annex A

THE LOGICAL FRAMEWORK

The logical framework is a highly effective planning tool for defining inputs, outputs, timetables, success assumptions and performance indicators. It provides a structure for specifying the components of an activity and for relating them to one another. It also helps to identify the place of a project within an overall program or a national research system. The table below illustrates how a logical framework can be used for a research program.

Table 1: Example of Logical Framework: Research Program

	Narrative summary	Verifiable indicators	Means of verification	Important assumptions
Then	GOAL contribution to overall economic development goals	processing and marketing labour opportunity income	national statistics	political stability no drastic environmental changes no change in world prices
If then	PURPOSE contribution to agricultural development purposes	increased production or better efficiency	adoption rate farm surveys	necessary services available economic environmental changes
If Then	OUTPUT new or improved product (variety) or technology	specifications on desired characteristics	laboratory stations records on-farm testing results	continuous support (budget, staff) underlying scientific reasoning correct no drastic environmental changes
If	INPUT actions to be taken under the program	work plan schedule of activities	quarterly reports accounting and other administrative records	budget disbursed as planned staff available

The first column of the matrix provides a narrative summary for the four levels of objectives for the program, namely:

Inputs: These are activities undertaken under the program, and the means (resources, staff) used to undertake them, with the expectation that implementing these inputs will lead to the production of the designated outputs.

Outputs: represent those achievements (variety, technology, knowledge) which derive from the inputs and are not dependent upon other activities. For example, a breeding program (input) can be implemented to develop a new variety of wheat (output) with some specific characteristics.

Purpose: this describes a desired agricultural development scenario for which the research output is necessary but not always sufficient. In the example of a breeding program, it is expected that if a variety with specified characteristics is available (output achieved), then farmers can get a third crop in after their summer rice, thus increasing food production (purpose achieved).

Goal: in the broad context of national development, this is a desired economic achievement for which the agricultural development purpose is necessary but not always sufficient. Here, the expectation is that if more food is produced in-country (purpose achieved), then fewer imports will be necessary, with positive effects on food security and the balance of payments (goal). The purpose, and especially the goal level, will be common to several complementary programs.

An important process in the logframe procedure is to ascertain whether the problem addressed at the purpose level can really be solved by research. If production is not increasing because the price and marketing mechanisms are not favorable to producers, the problem is unlikely to be resolved by biological research. Similarly, the goal of reducing foreign exchange costs for food imports could, in theory, be reached by curbing population growth to match food production. Evaluators need to be able to clearly articulate the cause-effect relationship which was presumed when a given objective was assigned to research.

The second column, "verifiable indicators", specifies what type of evidence could be taken as a sign of achievement at each level of objectives, and the third column, "means of verification", indicates how that evidence could be found and measured. Decisions concerning indicators of achievement and means of verification have consequences at two levels: (1) they set up ahead of time the standards against which actual results should be measured, and (2) they set up the blueprint for data collection and reporting during the implementation of the activity. If the means of achievement are correctly selected and the data are actually collected during implementation, future evaluations will be greatly facilitated.

The last column, "important assumptions", is too often taken for granted in development activities, and lists those factors not controlled by the program but which influence its implementation and chances of success. For example, changes in the world price of wheat could influence the purpose-to-goal relationship by making wheat production less interesting to farmers, even if a suitable variety is available. Assumptions are particularly important for the program at the input and output levels.

Inputs include not only the proposed activities included in a breeding program, but also the means necessary to implement these activities. The availability and timeliness of necessary funding, manpower, facilities and equipment may be determined by development and administrative agencies outside the control of the research leaders. If these agencies fail to deliver on schedule, the potential benefits of the research may be lost through no fault whatsoever of the research agency.

Annex B

INDICATIVE COSTS FOR AN EVALUATION

In general terms, current per diem rates usually lie in the range of US\$ 50 - US\$ 120. National agricultural consultants in developing countries receive fees at between US\$ 50 and US\$ 200 daily and international agricultural consultants US\$ 150 - US\$ 300.

From this starting point a hypothetical evaluation review has been costed for a developing-country agricultural program or institute. The review is scheduled to last three weeks and the review team to be composed of two persons from the national agency, one of whom acts as coordinator, two consultants from local universities and three overseas consultants. The table below gives costs based on these figures. It assumes that the two national-program scientists only receive the per diem rate for time spent in the field, and that the team leader spends one week at his home base finalizing the review report at the end of the mission.

Table 2: Indicative Costs for an Evaluation using National and Foreign Consultants

		US\$	%
<u>PERSONNEL</u>			
1 international consultant	28 days at US\$ 250	7000	16
2 international consultant	21 days at US\$ 250	10500	24
2 national consultant	21 days at US\$ 100	4200	10
<u>PER DIEMS</u>			
5 consultant	21 days at US\$ 70	7350	17
2 staff	14 days at US\$ 70	1960	4
<u>TRAVEL</u>			
3 international tickets	at US\$ 2000	6000	14
7 local tickets	at US\$ 500	3500	8
<u>REPORTING</u>			
Report preparation		3490	8
		44000	100

The total cost of the review is US\$ 44,000, a figure that is consistent with a number of reviews with which ISNAR has been associated. In looking at the detailed cost structure it can be seen that 40% of the total costs are fees for external consultants. If their air fares are included this figure rises to 57%. In contrast to this, consultancy fees for the four national consultants represent less than 10% of the total cost. Were all of the consultants to be nationals, the team could probably be reduced to 5 or 6 persons and the total cost would be about half of the figure cited.

Obviously there is an infinite number of combinations that can be used to build up a table such as the one above, depending upon the manpower composition of the team and the unit costs. The important point to consider, however, is that evaluations using expatriate international specialists can be expensive undertakings. This suggests that they should not be done too frequently and that particular care should be used in the selection of the right personnel for conducting this task if it is to be cost effective.

Annex C

EXAMPLES OF BACKGROUND TABLES

The list below illustrates the type of tabular material that may be useful for the evaluation review of a research institute or program. The list is intended to be illustrative rather than comprehensive, since each institute/program will call for specific information, some of which may be unique to that institute/program. However, a comprehensive evaluation will require information on most of the topics listed below. Much of this information is factual and can be assembled prior to the start of the evaluation process.

- A. Background Charts/Map
  - 1. Organization structure of research in the country
  - 2. Organizational structure of the unit being evaluated
  - 3. Map showing location of principal sites at which the unit operates
  
- B. Macro Data
  - 1. Area, yield and production of principal crops worked on
  - 2. Past trends in production
  - 3. Farm sizes in main production areas
  - 4. Current national plan targets (production and export) for main research commodities worked on
  - 5. Import prices and volumes
  - 6. Export prices and volumes
  - 7. Consumption data
  - 8. Price data
  - 9. Total research budget and budget for institute/commodity
  
- C. Institute/Program Data
  - 1. Land resources
  - 2. Physical Plant resources
  - 3. Budget sources, uses, trends
  - 4. Human resources, discipline, training level, numbers (including support staff)
  - 5. Training plans and targets
  - 6. Journals subscribed to by library
  
- D. Research Program
  - 1. Statement of mandate
  - 2. Precis of strategic plan
  - 3. Table of main research activities
  - 4. List of ongoing experiments (with budgets)
  - 5. Example of Research Protocol Form (completed)
  - 6. Examples of any monitoring or reporting form used
  - 7. List of varieties released or other innovations in last 10 years and extent of adoption
  - 8. List of publications over last 5 years and where published

Annex D

EXAMPLES OF REPORT OUTLINES

This annex presents the report outlines from two reviews with which ISNAR has been associated. In both cases checklists were used to draw up a report outline which was given to the evaluators along with their terms of reference. The latter included an instruction that they were expected to produce a report 'broadly in line with the report outline offered to them as a guideline'. This was not intended to be restrictive but to ensure that the evaluators knew precisely which issues the client agency sponsoring the evaluation wanted them to address. In this sense the checklist which forms the report outline is intended as a device to relate the final output to the objectives of the evaluation.

Example 1: Report outline for the evaluation of the program of a research institute XXXX in country AAAA

1. BACKGROUND

What is the role of agriculture in the economy of country AAAA

What are the agricultural goals of the latest 5-year plan

What is the status and role of the research institute

What are the physical resources available to the program:

- location of stations and farms
- status of lands and equipment

What are the human resources available to the program:

- existing numbers and level of training
- number currently in training, wastage
- training targets for the year 19--

What are the financial resources available to the program:

- current budget
- domestic/foreign components
- past and future trends

The specific objectives of the program research

- historic
- current

2. PLANNING AND PROGRAM FORMULATION

- PRIORITY SETTING:
- What are the priorities for research and the mechanisms for determining them? What kind of consultation (with policymakers, private sector, extension services and farmers) is involved in this process?

- What criteria are used in setting priorities? To what extent do these take account of available resources (human, physical and financial?) Do they take into account information based on economic analysis of empirical models?

- PROGRAM FORMULATION:
- How and who decides what research activities (projects) will be carried out in each priority area?
  - How does the program relate to the resources available to the institute? Is there an appropriate "program of work and budget" linkage?
  - How are the individual units within the institute coordinated in multi-disciplinary research projects?
- RELEVANCE OR THE PROGRAM IN TERMS OF THE OBJECTIVE:
- How does the current program relate to the objectives and priorities? Are there objectives not covered by the current research program?

### 3. IMPLEMENTATION

- PROGRAM ACTIVITIES:
- What research is being done? What is the main commodity, ecological or disciplinary emphasis? Is the research relevant to the goals of agricultural development? Is there duplication elsewhere in country AAAA? How does the physical location of the current research program related to both agricultural development requirements and resources available for research?
- METHODOLOGY:
- Is the methodology being used appropriate in relation to the research objectives and to the resources available?
- RESOURCE AVAILABILITY:
- Are the total resources (physical, financial and human) adequate for a realistic research program to be carried out on a continuing basis?
  - Is the manpower pool available appropriate for the research program in terms of experience level, disciplinary expertise and support staff mix?
  - Are funds available on schedule?
  - To what extent is the present program execution dependent upon external financing?
- RESOURCE UTILIZATION:
- To what extent does the research program effectively utilize the funds, lands and manpower available to it each year?

- Are the facilities and equipment being well maintained?
- To what extent does the program sub-contract research and with what results?
- Are there any formal monitoring and evaluation mechanisms being used by the program with respect to progress, expenditure control and staff assessment?

MONITORING AND  
EVALUATION:

4. LINKAGES

- WITHIN INSTITUTE XXXX: - What channels of communication exist-
- (a) within the program?
  - (b) with scientists in other XXXX programs?
  - (c) and how effective are they in both directions?
- WITH OTHER RESEARCH ORGANIZATIONS: - What channels of communication exist with-
- (a) other research institutes in country AAAA?
  - (b) universities?
  - (c) private sector?
  - (d) international institutes?
  - (e) and how effective are they in both directions?
- WITH EXTENSION SERVICES AND FARMERS: - What channels of communication exist with
- (a) field extension workers?
  - (b) farmers?
  - (c) and how effective are they in both directions?
- WITH POLICY MAKERS: - What channels of communication exist with
- (a) Institute XXXX headquarters?
  - (b) extension services?
  - (c) regulatory agencies?
  - (d) and how effective are they in both directions?

5. IMPACT

- RESULTS TO DATE: - How much of the program undertaken during the past 3 years has yielded tangible results?
- What are these results?
  - Do the results meet farmers' needs or are they likely to when the research is more advanced?
- USE OF RESULTS: - To what degree have the results been adopted by farmers?
- What are the main constraints to adoption?
  - Is research being undertaken to overcome these constraints?

6. CONCLUSIONS AND RECOMMENDATIONS

- A. PROGRAM PLANNING
- B. RESEARCH ACTIVITIES
- C. RESOURCE UTILIZATION
- D. LINKAGES
- E. COOPERATIVE ACTIVITIES

Example 2: Report outline for the evaluation of the commodity research program kkkk in the national research agency yyyy of country BBBB

## EXECUTIVE SUMMARY

### 1. INTRODUCTION

What were the objectives of the review?

What were the terms of reference (general and specific)?

Who were the personnel involved?

What was their itinerary?

### 2. BACKGROUND

What is the role of agriculture in country BBBB?

What is the role of commodity KKKK in the agricultural economy?

What are the major constraints in the production of KKKK and what quantitative effects do they have on production?

Research on KKKK in BBBB - What is being done to overcome the constraints:  
- by the national Agricultural Research Agency YYYY?  
- by others?

What agency YYYY resources are applied to commodity KKKK research?  
- human(no., balance, skills, leadership)?  
- physical?  
- financial??

(Section on resources should note manpower numbers, trend, level of education and should describe inputs from donors).

### 3. PLANNING AND PROGRAM FORMULATION

Priorities for commodity KKKK Research

- Who sets them?
- How is it done?
- Who is consulted?
- What criteria are used?
- Is the resource base considered?

How are commodity KKKK priorities related to the total research picture?

What are the current priorities?

- How much of the KKKK budget goes to them?
- Have priorities changed in the last 5 years?
- If so why?

What measures are taken to plan for the future; i.e., new diseases or shifts in commodity emphasis?

Program formulation for commodity KKKK research

- Who does it?
- How?
- Who is consulted, are changes made as a result?

How much awareness is there of work done elsewhere on similar problems?

How much use is made of literature searches?

How does the program which results relate to

- the priorities?
- research carried out by organizations other than YYYY?
- the resources available?

Is an ex-ante evaluation made of the potential benefits?

Is (large/small) farm size considered in program formulation?

If more funds were available where would they be spent?

If less funds were available where would cuts be made?

What % of the agreed annual program is actually carried out?

#### 4. RESOURCES

Are current resources adequate for an appropriate national research program on KKKK?

Are resources sufficiently stable from year to year?

Does resource availability in any area depend unduly on non-YYYY funding?

If so does this affect the nature of the program?

Is resource use internally consistent (can the available resources do the job)?

Is resource use externally consistent (is it doing what producers want)?

Are there constraints in the use of existing resources?

- physical (lack of spare parts)?
- financial (cash flow)?
- personnel (turnover, training)?

Are current programs to overcome these constraints appropriate?

To what extent are available resources utilized effectively (i.e. are funds unspent, is land underutilized, is the work load per scientist appropriate)?

To what extent are resources conserved by using joint programs or subcontracts with universities or other research organizations?

## 5. PROGRAM ACTIVITIES

What is the current research program - are its main thrusts commodity, disciplinary or ecological?

Does the program address the problems, is it likely to provide useful solutions?

Does it duplicate work elsewhere in country BBBB?

Is there an appropriate balance between basic and applied/adaptive research?

Is the location of the research appropriate?

Do funds constrain off-station trials and visits?

How well is the field work managed?

Is the methodology appropriate to the objectives and resources?

How well do project protocols cover an appropriate literature search?

Is an appropriate progress reporting system being used?

Is the final reporting appropriate and used? Are results analysed promptly and completely?

What sort of end evaluation takes place?

- how effective is it?
- what does it evaluate?
- what sort of economic criteria are used?

Are there any major gaps or deficiencies in the program?

## 6. COMMUNICATION (LINKAGES)

Does the program have appropriate linkages with policymakers?

Does the program have appropriate linkages with the non-YYYY scientific community in country BBBB and overseas (CGIAR, universities)?

Does the program have appropriate linkages with related programs/disciplines in YYYY?

Does the program have appropriate linkages with extension personnel and farmers?

How effective are these various linkages in BOTH directions?

## 7. IMPACT

Results to date

- How much of the program in the last 5 years has given tangible results?
- What are they?
- Do they or will they meet the farmers' needs?
- What type of farmer will be the main beneficiary?

- Use of results
- What has been the extent of adoption?
  - What are the constraints to adoption?
  - What research is being done to overcome these constraints?

## 8. CONCLUSIONS AND RECOMMENDATIONS

Overall conclusions:

Specific recommendations with respect to:

- planning and program formulation
- resources and their use
- research activities
- communications
- potential areas of cooperation with other organizations (public, private, international).