Training in Planning, Monitoring and Evaluation for Agricultural Research Management

Manual 4
Evaluation

Alicia Granger
John Grierson
Tarcizio R. Quirino
Luis Romano

nar
National Agricultural Research

1995

CIAT
International Center for Tropical Agriculture
Centro Internacional de Agricultura Tropical

Donors

Funds for preparing and publishing the series of training modules and manuals "Training in Planning, Monitoring and Evaluation for Agricultural Research Management" have been provided by ISNAR's core budget with a special contribution from the Government of Spain and the following donors:

- The Inter-American Development Bank (IDB)
- The International Development Research Centre (IDRC)
- Swiss Development Cooperation (SDC)
- Technical Centre for Agricultural and Rural Cooperation ACP-CEE (CTA)

Correct citation

A. Bojanic, G. Hareau, R. Posada, A.M. Ruíz, and E. Solís. 1995. Monitoring in Agricultural Research Management. Manual 3 of the series *Training in Planning, Monitoring and Evaluation for Agricultural Research Management.* The Hague: International Service for National Agricultural Research.

Translated from A. Bojanic, G. Hareau, R. Posada, A.M. Ruíz, and E. Solís. 1995. Seguimiento en la Administración de la Investigación Agropecuaria. Fascículo 3 de la serie *Capacitación en Planificación, Seguimiento y Evaluación para la Administración de la Investigación Agropecuaria*. La Haya: Servicio Internacional para la Investigación Agrícola Nacional.

Published in collaboration with the Technical Centre for Agricultural and Rural Cooperation ACP-CEE (CTA), The Netherlands; the Bolivian Institute for Agricultural Technology (IBTA), Bolivia; the National Institute for Agricultural Research (INIA), Uruguay; the National Institute for Agricultural Technology (INTA), Argentina; the Ministry of Agriculture and Livestock (MAG), Costa Rica; and the Train-the-Trainers Project (ISNAR-CIAT agreement).

Send us your ideas

Training materials such as these are not finished products but work in progress. They can always be improved. Since we hope to revise them in future, ***the authors and ISNAR would appreciate receiving your comments and suggestions for improving them***. We would also be interested in learning about your experiences (positive and negative!) using these materials in training and in institutional-change processes.

Manual 4
Evaluation in Agricultural Research Management

INTA
National Institute for Agricultural Technology, Argentina

INIA
National Institute for Agricultural Research, Uruguay

EMBRAPA
Brazilian Corporation for Agricultural Research, Brazil

ICA
Colombian Institute for Agriculture and Livestock, Colombia

Authors (in alphabetical order):
Alicia Granger
John Grierson
Tarcizio R. Quirino
Luis Romano

Coordinators:
Vicente Zapata S.
Douglas Horton
Juan Cheaz

Reviewers:
George Norton
Luis Zavaleta

Production:
Viviana Gonzalías

Design:
Juan Carlos Londoño

1995

isnar CTA CIAT

Content

Page

Preface 2
Acknowledgments 2
Flowchart for Manual 4 3
Introduction to Manual 4 3

Sequence 1. Frame of Reference for Evaluation 5
- Flowchart for Sequence 1 6
- Conceptual Elements of Evaluation 6
- Levels and Uses of Evaluation 9
- Objects of Evaluation 11
- Types of Evaluation 13
- Present Situation of Evaluation in Latin America and the Caribbean 17
- Summary 21

Sequence 2. Evaluation Methodology 23
- Flowchart for Sequence 2 24
- Evaluation Design 24
- Implementation of Evaluation 28
- Conclusion 37
- Summary 37

Sequence 3. Relating to other Institutional Processes 39
- Flowchart for Sequence 3 40
- Dissemination of Evaluation Results 40
- Institutionalizing Evaluation 44
- Conclusion 49
- Summary 49

Appendices 51
Appendix 1. Terms Used in the PM&E Manuals 52
Appendix 2. Summary of the Types of Evaluations Used in Research Organizations 56
Appendix 3. Individual and Collective Variables 57
Appendix 4. Evaluation Scale for Interviewers 58
Appendix 5. Planning for Computerized Analysis and Data Processing 58
Appendix 6. Planning, Monitoring and Evaluation in the National Institute for Agricultural Technology (INTA), Argentina 59
Appendix 7. Bibliography 66

Preface

During the ISNAR project "Strengthening Agricultural Research Management in Latin America and the Caribbean" a team of individuals representing national, regional, and international organizations produced several publications and training materials on planning, monitoring, and evaluation (PM&E) for agricultural research institutions in Latin America and the Caribbean.

These materials were designed to:

- support learning and training courses and workshops on PM&E;
- facilitate the diffusion of concepts, methods and tools for improving PM&E in the region and elsewhere.

Three types of materials were developed: reference books, training modules, and training manuals. The training *manuals* are intended for course and workshop *participants*; the training *modules* are to be used by *instructors*. In this sense, the manuals and modules are complementary. The manuals present the training objectives and essential subject matter. In the modules, these components are complemented with special sections for instructors, including exercises, transparencies, and technical annexes. Instructors and course participants who want additional information about the topics discussed in the materials can turn to the project's reference books or to the many references in the course material.

We hope that managers and trainers working in agricultural research will find these materials useful. We hope they will not only distribute them in their institutions but also apply the concepts and tools discussed.

Acknowledgments

The authors would like to express their thanks to the various individuals and institutions that made it possible to produce this training module on evaluation in agricultural research management.

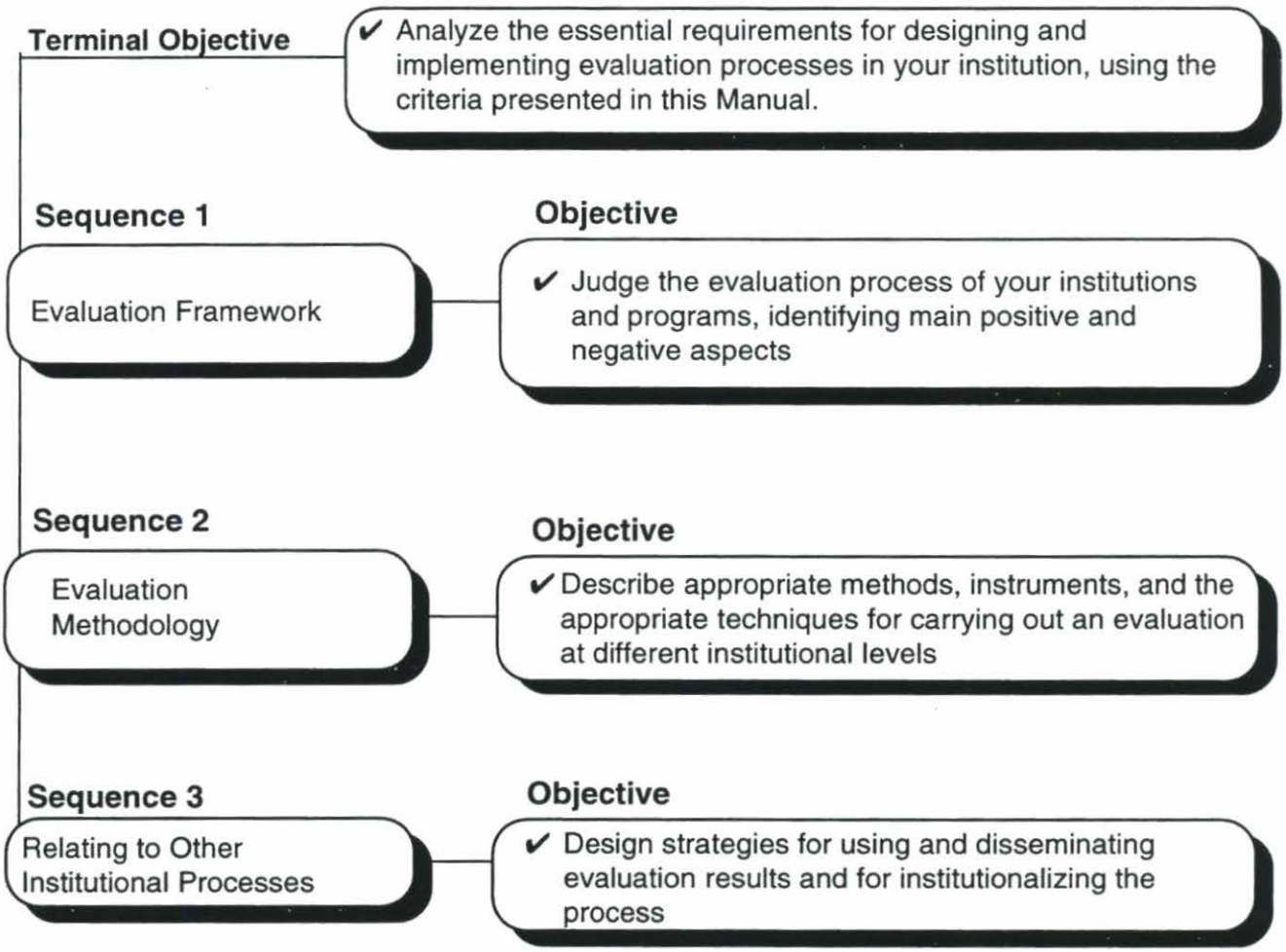
We would like to thank Christian Bonte-Friedheim and Gustavo Nores, Directors General of ISNAR and CIAT, for establishing the inter-center collaboration that facilitated the production of these training materials. We would like to give special thanks to Douglas Horton of ISNAR for his leadership in the project and for giving us the opportunity to participate in it. We are indebted to Juan Cheaz for the efficient arrangements he made for project events and for his dedication and many contributions during preparation of the training materials.

We are thankful to the Inter-American Development Bank (IDB), the International Development Research Centre (IDRC), the Swiss Development Cooperation (SDC), the Technical Centre for Agricultural and Rural Cooperation (CTA), the Government of Spain, and ISNAR for providing the necessary funds for preparing and publishing this module.

We would like to thank Gerardo Häbich, Associate Director for Institutional Relations of CIAT, for the support and hospitality that he arranged for us at CIAT. Vicente Zapata, Train-the-Trainers Project Coordinator, and the entire team of CIAT's Training Materials Unit guided and supported us in our development as instructors and in preparing this training module. The skill and patience of the CIAT team throughout the numerous revisions of this module are much appreciated. In particular, Viviana Gonzalías demonstrated a high degree of professionalism and dedication and helped us throughout the complex and tedious process of preparing the text and supporting materials. We would also like to thank Flora Stella de Lozada, who ably transcribed the initial materials; Juan Carlos Londoño for his numerous and invaluable contributions to the design and production of the final module and transparencies.

Finally, we would like to express our gratitude to our own institutions, which kindly relieved us from our normal duties to allow us to participate in the various activities in this project, to develop our training skills, and to prepare these training materials.

Flowchart for Manual 4



Introduction

This is Manual 4 in the series on training in Planning, Monitoring, and Evaluation for Agricultural Research Management.

In general, this manual should follow the three preceding manuals on the Strategic Approach, Strategic Planning, and Monitoring Agricultural Research , respectively. It can also be used for specific training needs in evaluation; in this case, the instructor could summarize the basic material covered in the previous manuals by quickly reviewing their contents before starting on this one. Information in this manual is grouped in three Sequences or chapters, which contain the following material: Sequence 1 provides a general framework for evaluation. Evaluation, and its purpose, levels and types are defined. This sequence also describes the present situation of agricultural research evaluation in Latin America and the Caribbean.

Sequence 2 deals more specifically with evaluation methodology, including activities for designing and implementing evaluations. Data collection, analysis, and interpretation are examined within the framework of agricultural research evaluation. Finally, Sequence 3 treats management and administrative aspects of evaluation in more detail.

It includes techniques for communicating evaluation results, taking into account such aspects as periodicity, institutional structure, and the necessary technical team.

The logic implicit in this manual is organized so that Sequence 1, which is conceptual, attempts to answer the question "Why carry out an evaluation"? Sequence 2, which is methodological, clarifies "How to carry out the evaluation"?. Sequence 3, which is strategic, answers the questions, "Who should carry out the evaluation"?, and "For whom should it be done"? These process components, in terms of Context, Inputs, Processes, and Products of the evaluation, are illustrated in Figure 1.

With this information, participants should be able to apply and reinforce the processes of research evaluation in their own agricultural research institutions.

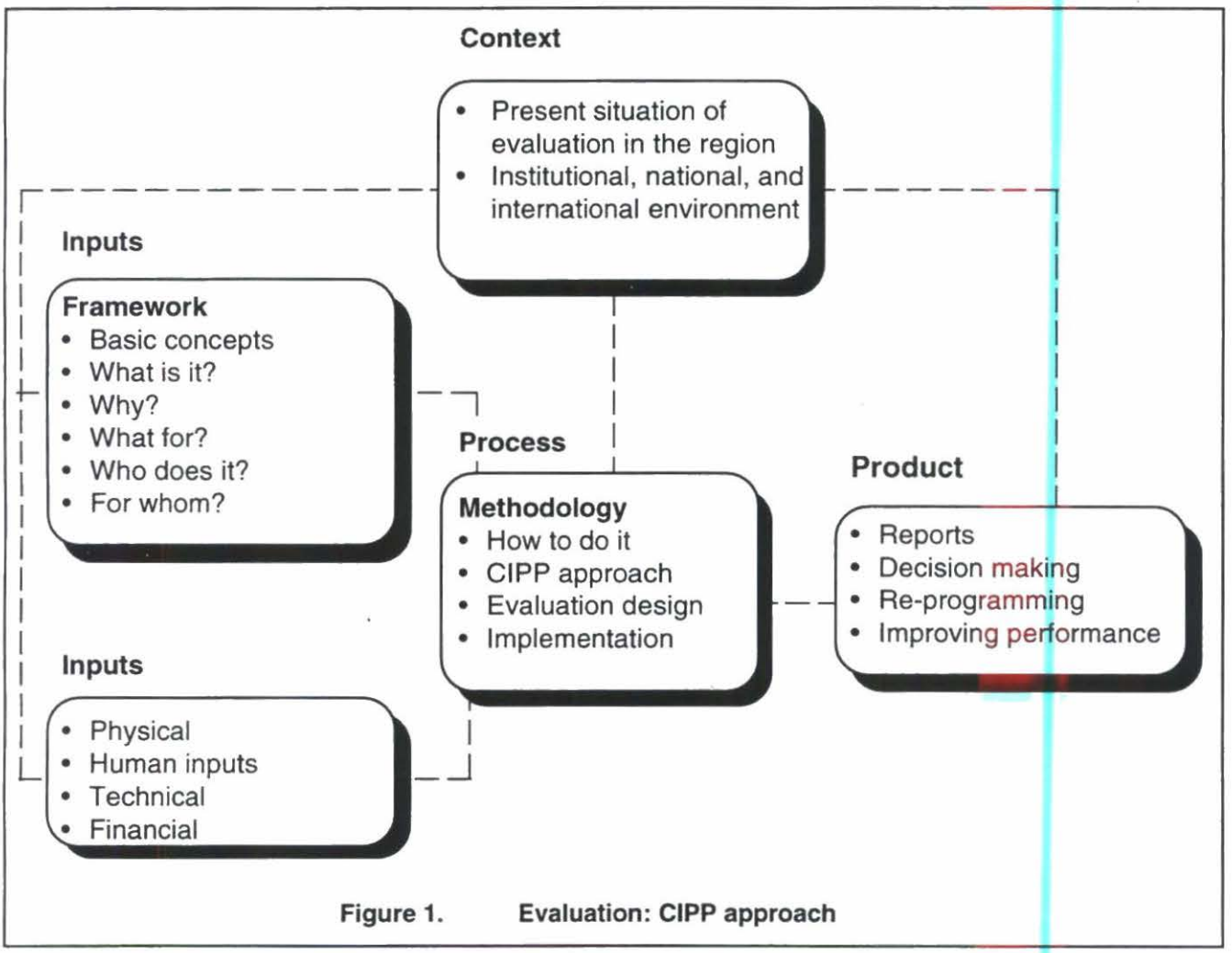


Figure 1. Evaluation: CIPP approach

Sequence 1. Frame of Reference for Evaluation

Page

Flowchart for Sequence 1 6

Conceptual Elements of Evaluation 6

Levels and Uses of Evaluation 6

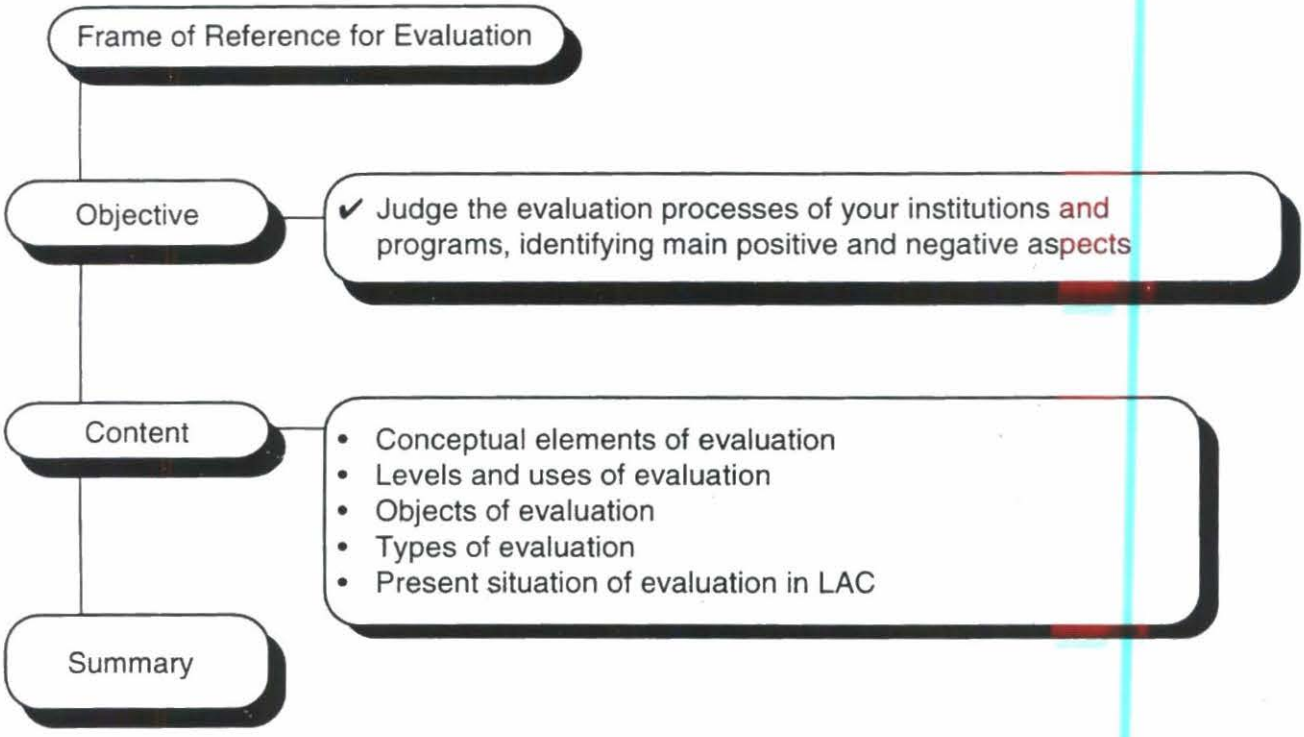
Objects of Evaluation 9

Types of Evaluation 11
- *Ex-ante*, on-going, and *ex-post* evaluation 13
- Internal and external evaluation 13
- Organization of internal and external evaluations 14
- Other types of evaluation 14

Present Situation of Evaluation in Latin America and the Caribbean 17
- Experiences and methods 17

Summary 21

Flowchart for Sequence 1



Conceptual Elements of Evaluation

As previously recognized (Manual 1), research institutions operate in a national and international environment whose social, economic, political, and technological aspects change frequently. This implies not only opportunities for progress but threats to institutional sustainability. Evaluation can provide a means to judge the value or merit of different areas or aspects of research, (for example, proposals, on going experiments, or completed projects) in relation to changing external conditions.

Evaluation has always been part of the tasks of agricultural scientists; but systematic use is recent and not generalized. This is also true of the need for training in managerial techniques. Scientists who are highly qualified for carrying out research are often also responsible for leadership. They are placed in administrative positions without any training in management or the administration of technological development. They need to be provided with the basic elements for exercising their new functions, which involve managing the processes of planning, organization, direction, and evaluation (Fonseca, 1990).

Evaluation, as a tool of research management, can be used by researchers and managers to improve institutional performance and agricultural technology and to gain support for the research process. The term evaluation has diverse meanings. It has been understood as the basic review, control, and supervision of processes and activities, but also as the more complex socio-economic studies on the impact of research results.

A comprehensive definition of evaluation, including the essential elements of the process follows:

> **Evaluation is judging, appraising, or determining the worth, value, or quality of research —whether it is proposed, ongoing or completed— in terms of its relevance, effectiveness, efficiency, and impact.**

Relevance indicates how appropriate or important the goals and objectives are with respect to the established needs. **Effectiveness** measures the

degree to which the goals have been reached. **Efficiency** refers to the cost effectiveness of an activity. And, **impact** to the broad and long-range effects of research (Horton *et al.*, 1993).

Evaluation is also part of the organizational process for improving on-going activities and backing of future planning and decision-making activities. Evaluation should not be considered as an isolated activity, but as part of the cycle that begins with planning, and monitoring, is followed by evaluation, and ends with decision leading to a the new management cycle. (Figure 2).

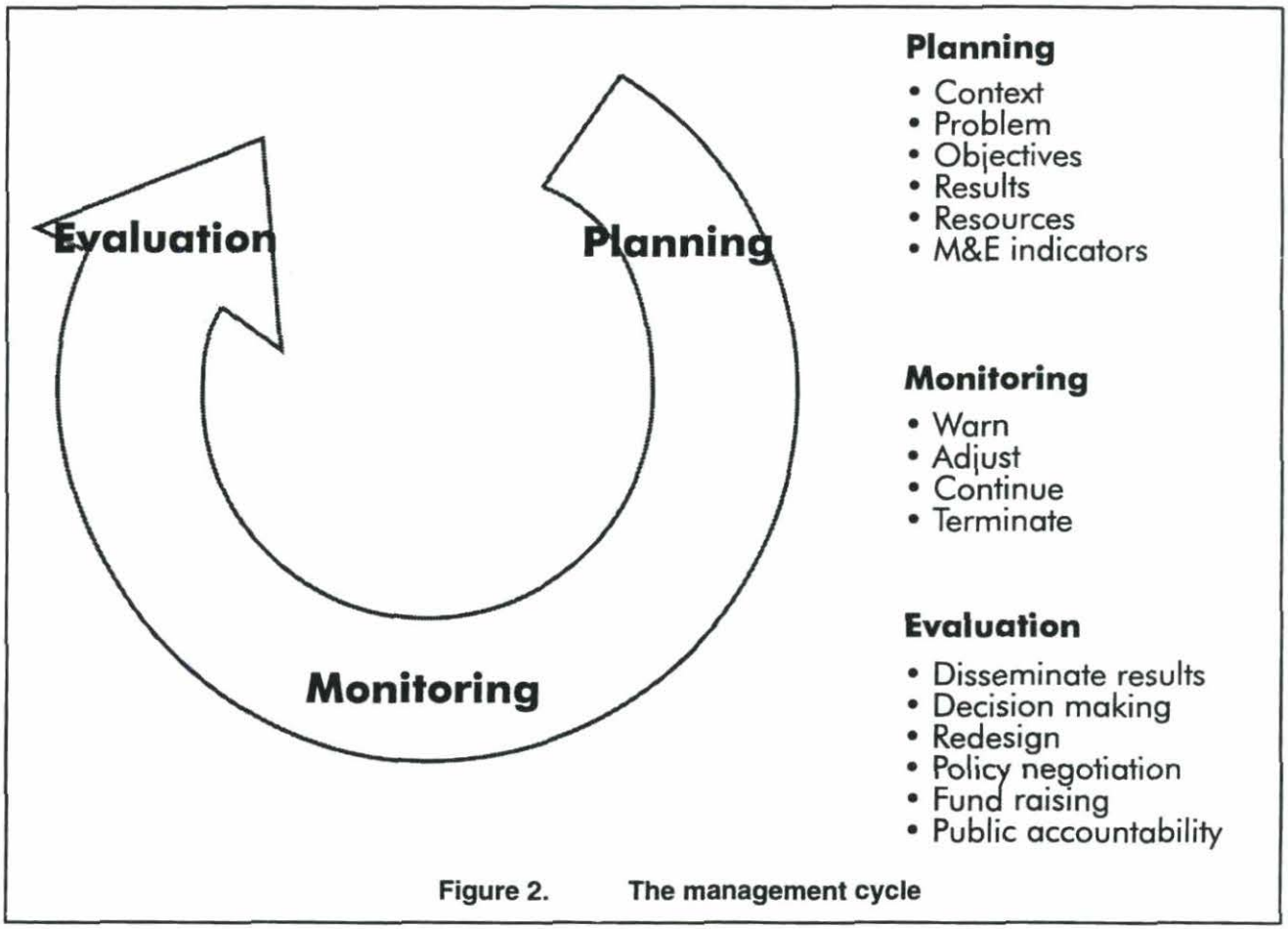


Figure 2. The management cycle

Murphy (1993) emphasizes the following principles of evaluation:

Principle 1. An evaluation begins with the design of an activity with clear objectives and verifiable indicators of the achievement of those objectives.

An evaluation always involves a relative judgment. One can only judge a situation by comparing it with another. For this reason it is necessary to arrive at an agreement on what will be acceptable as an indicator or measure of achievement.

These comparisons require detailed and credible information. Consequently, the first rule of evaluation is that it begins when an activity is designed —defining clear objectives for it, selecting specific, verifiable indicators of achievement— and stating how the achievements will be measured. This requirement gives us two immediate benefits, which go beyond the evaluation:

- It makes program designers express the objectives and expected results clearly and specifically. This may generate considerable discussion among planners and researchers, since individuals may often agree upon the need for a program, but disagree as to its objectives.

- It requires a specification of how progress and achievements will be measured, thus establishing the bases for monitoring. This brings us to the second key principle.

Methodologies are available to facilitate design and specification tasks, such as the logical framework, presented in Manual 3.

Principle 2. **Evaluations are more effective if adequate monitoring, recording, and information mechanisms are actively implemented during the course of the program.**

In order to compare results achieved with those expected, it is indispensable for the evaluating group to have up-to-the-minute data that indicate what has really been achieved, and what has happened during the process. It is difficult, time-consuming, and often impossible to obtain this information a posteriori. It is much easier to follow the development implementation of the program through systematic records and periodic reports. But this internal collection of data is not enough, as the third principle indicates.

Principle 3. **Research activities form part of a research and development system; therefore, evaluation has to situate the activity in the institutional, political, social, and economic context in which it is carried out.**

Knowing whether the expected results were achieved or not, is not enough to judge the quality of a program's design and execution. It is also necessary to understand why some results were achieved and others not, and to distinguishing clearly between the internal factors of the program and those external to it.

There are numerous examples of programs correctly designed and competently executed that did not attain the expected results due to external factors, such as changes in key personnel or budget reductions. When an evaluation discovers that the expected results, or some of them, could not be achieved, it is essential to establish whether this was due to restrictions beyond the control of researchers. This is not intended to provide the program leader with an easy excuse, but rather to identify the real causes of the problem and, from there, define potential solutions.

Principle 4. **Evaluation of a research activity must clearly distinguish between the achievement of research objectives and the contributions of these results to broader development objectives.**

Both appraisals of evaluation are valid, and in fact, complementary, but not interchangeable. An evaluation that is specifically limited to satisfactorily fulfilling an objective includes internal factors (design, execution, etc.) and external factors that have influence over resources, implementation, scientific soundness, and the potential of the results. Results evaluated are those derived directly from the activity, such as a new crop variety or a set of improved practices.

An evaluation of a program's contribution to a broader development objective will cover the same factors, but will add two more elements: (a) how the program fits into the overall national research and development plan, and (b) whether program results were adopted, and what their impact was on production, on the income of producers involved and on any other objective of the development plan. This introduces numerous factors not related to scientific research-factors that are regional, national, and occasionally international.

It may be necessary to analyze other research programs for a better understanding of the contribution made by a specific program to the development of the sector.

This gives rise to two conclusions. First, that evaluating the contribution of research to development is more complex than evaluating what has been carried out, because it includes a much wider range of aspects and requires different know-how and expertise. Second, its results have broader applications in overall planning and in establishing priorities, not just for research but also for services more directly related to development.

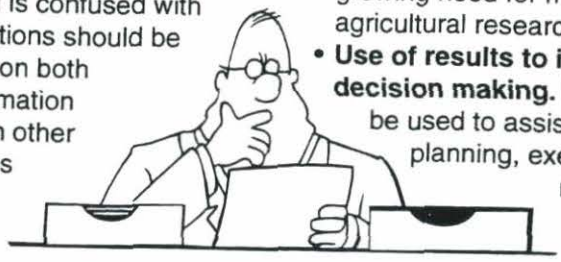
Evaluating the impact on development of a set of research activities will be much easier if each individual activity is evaluated first, but only if these more limited evaluations are carried out using comparable procedures. This does not mean that

the first evaluations must be carried out by the same team of evaluators, or that the evaluations must have an identical scope; co-ordinating the preparatory stage is the essential element.

Evaluation principles

- An evaluation begins by designing the activity, and this, in turn, begins by defining clear objectives and selecting specific and verifiable indicators of the achievement of objectives.
- Evaluations are more effective if adequate monitoring, recording, and information mechanisms are actively implemented during the course of the activity.
- Research activities form part of a research and development system; therefore, evaluation has to situate the activity in the institutional, political, social, and economic context in which it is carried out.
- Evaluation of a research activity must distinguish between the achievement of research objectives and the contributions of these results to broader development objectives.

Sometimes, on-going evaluation is confused with monitoring. The following distinctions should be considered: evaluation is based on both quantitative and qualitative information compiled by monitoring and from other sources. While monitoring follows research achievements and establishes whether progress is being made in accordance with plans, evaluation analyzes problems of quality and relevance, and can even analyze the appropriateness of the plan itself. Monitoring and evaluation share some sources of information, as observed in Figure 3. In an integrated process they also share some mechanisms, such as internal reviews. Evaluations result in a set of recommendations directed towards **planning**, such as a change of objectives, or towards **implementation**, to improve research design, or decision making, for example allocating new resources.



Levels and Uses of Evaluation

Evaluations can be carried out at the different decision-making program levels within agricultural research institutions, from the level of an experiment, project or program, to the level of a national research institution or system (set of institutions and resources) (Figure 4).

The results of an evaluation can be utilized with different purposes at different management levels, in different stages of research. The users of evaluation results may include managers and decision-makers, policymakers concerned with research, government authorities, research program coordinators, center directors, and social groups with an interest in the research benefits. The Section "Uses of evaluation results" in Sequence 3 of this module expands on this point.
When carried out independently and objectively, the results of both external and internal reviews can influence decisions made by the top management levels on organizing or strengthening research.

Among the main uses of evaluation, two stand out:

- **Uses related to public accountability.** This refers to the responsibility of individuals and organizations to use resources properly. Traditionally, this responsibility has been carried out by means of periodic reports on the activity and the use of resources. However, there is a growing need for more and clearer evidence of agricultural research results and impact.
- **Use of results to improve management and decision making.** Evaluation can also be also be used to assist decision making during the planning, execution, and the periodic review of research activities. Accountability and decision making should be closely related. For example, the information supplied by a scientist or a research organization to meet public accountability requirements can also be used by top management to argue for future financial sources for research.

Accountability is also a part of the good management of an organization. Senior managers must ensure that their subordinates and those managing the projects can answer for the resources they use. Social groups in general are increasingly concerned with how their tax money is being spent. Applied research is controlled to ensure the inputs, activities, and results are directed towards the objective established, in order that the management can correct any deviations from that goal. This theme is covered more fully in the section "Uses of evaluation results" in Sequence 3 of this manual.

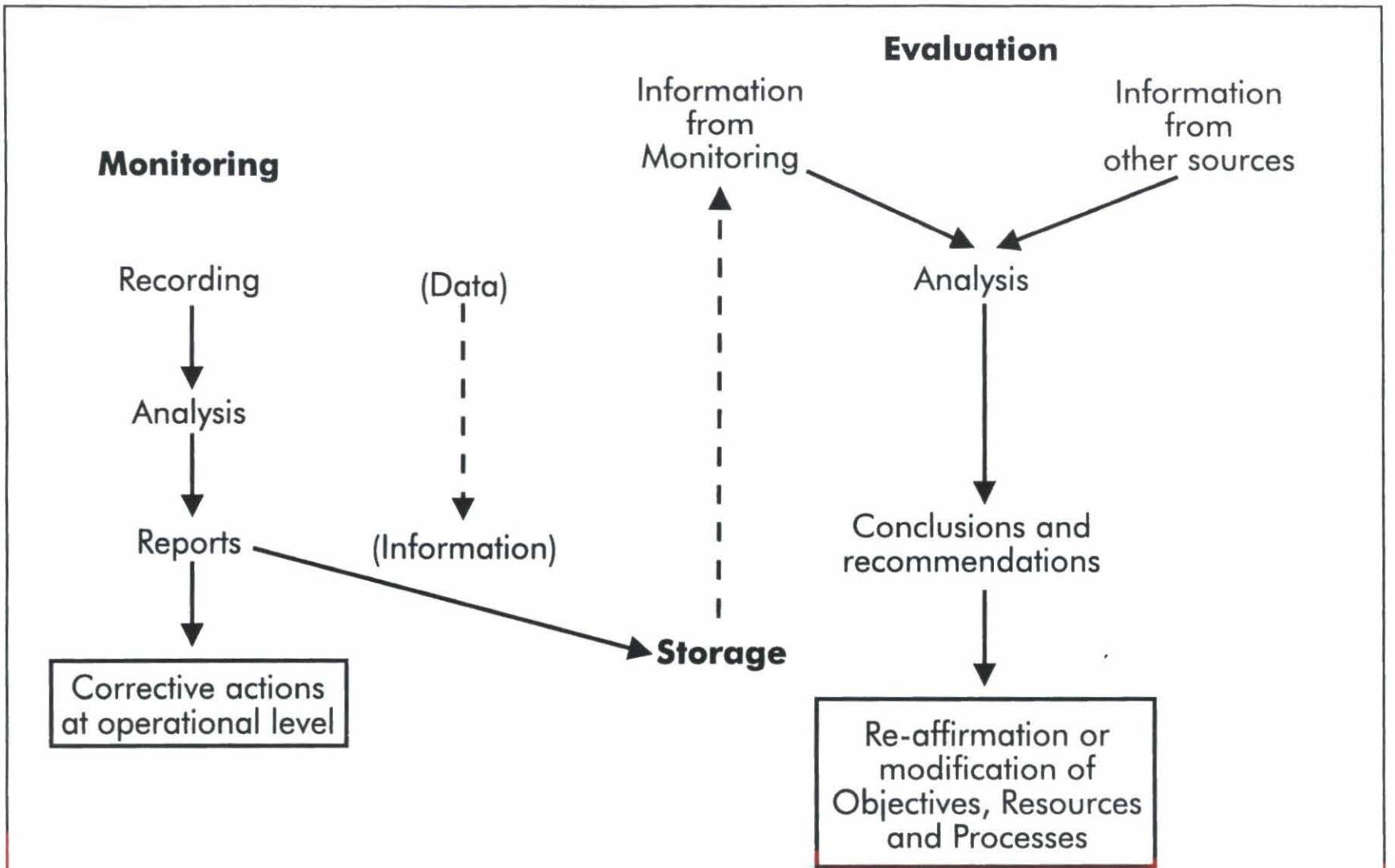


Figure 3. Relation between monitoring and evaluation

Source: McLean, 1988. Monitoring and Evaluation in the Management of Agricultural Research. Working Paper 14. ISNAR: The Hague

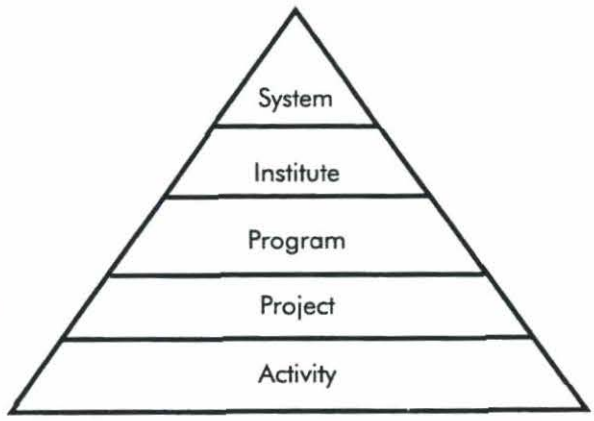


Figure 4. Levels at which evaluation can be carried out

Objects of Evaluation

Many different things can be evaluated within the scope of agricultural research: activities, projects and programs; research resources (e.g. scientists, funds, and physical contributions); research organizations (experiment stations, laboratories, or institutes); national research systems made up of diverse organizations; as well as research results (technologies and information) and their impact. The level (institute, program, project, etc.) and the components being evaluated at each level (inputs, processes, products, context) need to be defined and distinguished.

Defining precisely what is to be evaluated is more difficult than one would think, because the object of the evaluation is often tied up with other elements of the research system. For example, it is difficult to separate the management of a livestock program from the management of the research station in which the program functions. It is also difficult to evaluate the effects of a project funded by donors that has been incorporated into a wider national program. When defining the object of evaluation, it is crucial to make note of its main purpose, assumptions, components, inputs, and products, as well as notes on the context within which it is found and operates.

Specifically, the objects of evaluation are the context, inputs, processes, and products present at each research level. These objects can be established as follows.

Context evaluation identifies the target population and its needs, diagnoses current problems, identifies opportunities, and judges whether the proposed objectives meet the needs. In general, it is used for selecting objectives and determining priorities. It includes changes in the social, economic, political, institutional, scientific, and technological environment.

Input evaluation refers to all inputs to the research process itself —including research strategies, design and procedures as well as the human, financial and physical resources.

Process evaluation identifies implementation processes and problems, providing information for improving implementation or redesigning procedures. It includes budgeting, management, research, and operational processes, among others.

Product evaluation describes and appraises results in relation to objectives, goals, and clients' needs. It seeks to measure the value and merit of the final result. The results of this evaluation provide guides for finishing, continuing, or modifying research activities.

When the evaluation appraises the effects of results on the broader environment (e.g. the impact of new technologies on yields, production or prices) it is called an **impact evaluation**.

Different methods of analysis may be needed for each level of analysis and object of evaluation. In the case of an impact evaluation, for example, socioeconomic methods are commonly used.

The components of evaluation —Context, Inputs, Processes, and Products (CIPP)—are illustrated in Figure 5.

As mentioned earlier, the components of the CIPP model (objects of evaluation) are present at all research levels. The interrelationship between levels and components is seen in the grid in Figure 6.

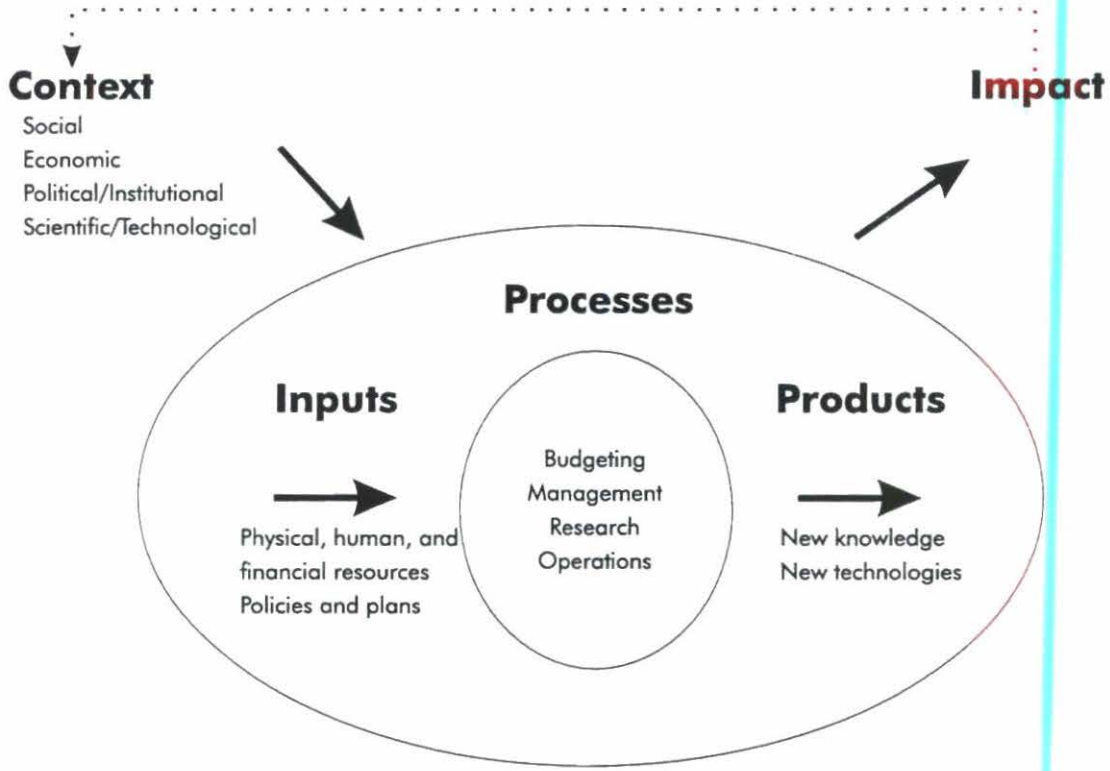


Figure 5. Objects of evaluation: CIPP model

Levels of research	Types of evaluation			
	Context	Inputs	Processes	Products
System				
Institution				
Program				
Project				
Activity				

Figure 6. Example of a grid showing the inter-relationship between management levels and types of evaluation

Types of Evaluation

Ex-ante, on-going, and ex-post evaluation

Depending on the moment when an evaluation is carried out, it can be classified as *ex-ante*, on-going, or *ex-post*.

Ex-ante evaluation

This is done before carrying out research and is used in the decision-making process in order to choose the best option or project from among several options. Various authors propose four methods for ex-ante evaluation:

- Scoring models.
- Cost-benefit analysis.
- Simulation.
- Mathematic programming.

Ex-ante evaluation is not practiced widely, although its value is recognized in:

- supporting the research planning process and helping leaders identify low-cost, priority alternatives for allocating resources;
- assessing the importance of projects and the soundness of proposed methodologies; and
- avoiding duplications.

Besides assessing technical possibilities and economic feasibility, *ex-ante* evaluations can also provide information on the social acceptance and the environmental consequences of proposed research outputs.

Ex-ante evaluations of projects may be carried out to establish which ones have the greatest chance of success.

On-going evaluation, or supervision

Evaluations carried out during implementation can lead to timely decisions and actions leading to the achievement of previously established goals.

Supervision can draw on an information system or on direct observations, which provide managers with information on the current state of the projects. It should indicate deviations, delays, failure to accomplish goals, and other problems requiring immediate attention.

On-going evaluation, or supervision, generally centers on the implementation process and is related to the following activities: ensuring objectives and methodologies are still valid; reviewing the efficiency of the process, or checking that results agree with plans; determining the adequacy of inputs; calculating the probability of success; giving the feedback required to apply corrective or complementary measures that improve the project's development; and encouraging communication and co-ordination among implementing units. This type of evaluation is similar to monitoring but this is more concerned with relevance and with quality as well as fulfilling objectives.

Ex-post evaluation

Ex-post evaluation checks whether the new knowledge and technologies produce correspond to initially-established goals or present needs. It constitutes a critical analysis of research achievements and results in relation to proposed objectives, basic assumptions and strategies, and the use of resources. *Ex-post* evaluation is a process by which the things that have been proposed can be compared to what has really been achieved, taking into account how, why, and with what it was done. It also identifies the reasons for satisfactory or unsatisfactory results.

The relation between the products of research and the inputs used can be expressed in cost-benefit coefficients. Relations between the products and the context of research can be expressed as effect or impacts.

Ex-post evaluation also produces results and pertinent lessons for future or on-going activities, and these should provide feedback for the system and become useful indicators for the *ex-ante* evaluation.

Ex-post evaluation can be used mainly for:

- Operational ends, to verify achievements
- As an analytic tool, to improve the design and methodologies of new proposals, or as feedback for on-going research
- Policy ends, to verify the validity of a strategy or approach.

The evaluation of finished projects is aimed at identifying: technologies generated needing more field trials and adjustments; technologies ready to be disseminated and used; new research areas needing attention; and information for establishing policies.

Internal and external evaluation

An integrated PM&E system allows for timely planning and adjustments of activities, and satisfies both external and internal demands. Such a system should include two types of evaluations: internal and external.

Internal evaluations

Internal evaluations are generally carried out at the level of the center (experiment station or institute), program, project, or technical area. Annual program or project reviews are organized with the participation of researchers, managers, technicians, and, at times, external stakeholders. These reviews, are designed to analyze results, achievements and breakthroughs as well as problems and limitations on achieving results. These internal evaluations can provide a critical analysis of activities and progress.

Internal evaluation permits the adjustment of activities for the following period and the collection of information for short-term decision making. It also favors the integration of researchers and processes, and stimulates or encourages researchers work.

These evaluations should be documented in reports or proceedings that supply data to other technical, management, and political levels of the system, mainly for short-term decision making.

External evaluation

External evaluations are carried out to review a research center, experiment station, program, or project, with the participation of external evaluators. These should be carried out every 3-5 years. In general, they are organized by top management and guided by institution's policies.

External evaluations include: the analysis of overall progress of the unit, taking into account the different demands; the priority of the activities; results and achievements; resources availables; and any difficulties or problems that can affect the operation of the institution. Evaluation provides new ideas which contribute to both proposals and to the results expected.

The result of an external evaluation should be presented to the institution's top management; the report may later be circulated to other institutional levels and groups involved.

Organization of internal and external evaluation

Table 1 contains some of the advantages and disadvantages of internal and external evaluators. Organizing an evaluation includes preparing the evaluation's objectives and terms of reference, selecting the evaluating group, writing and presenting the report, and monitoring the entire evaluation process.

Other types of evaluation

Depending on the objects being evaluated, the types of evaluation can be related to each institutional level (Table 2). The chart presents numerous combinations for the use of the various types of evaluation. McLean (1988) suggests several uses for evaluation information at each research management level, and the methods of analysis most frequently used (Table 3).

One important method of agricultural research evaluation is **peer review**. In this, scientific merit is evaluated by scientists familiar with the research area. The following are types of peer review:

- **Direct**, in which peers determine the scientific merit of an activity, usually in committees, both for ex-post and ex-ante evaluations.
- **Modified**, similar to direct evaluation except that criteria are broadened to cover the socioeconomic aspects of strategic or applied research by including non-scientific participants,
- **Indirect**, based on information from previous reviews carried out by peers and conducted for other purposes,

Another method used in agricultural research evaluation is **expert review**. It involves eminent specialists in a certain field and is generally used to evaluate complete programs or specialized institutes.

Other types of evaluations respond to needs regarding financial responsibility, demonstrating the economic impact of research, and staff performance.

Table 1. Advantages and disadvantages of internal and external evaluators

Advantages	Disadvantages
Internal evaluators	
• Familiarity with the program and how it operates • Consistency with the institute's values • Less time required for scheduling evaluations • Lower cost	• May not be objective • Possible conflict among organizational functions • Difficulty in freeing staff from their regular activities
External evaluators	
• More objective • Free of institutional bias • Greater possible access to decision makers • Time assigned to the task	• May be perceived as "alien" and make staff anxious • Require time to negotiate the contract • Higher cost
Collaborative evaluations	
• Have the advantages of both internal and external evaluators • Can promote learning and strengthen internal capacity for evaluation	• Open discussions on sensitive subjects may be limited

Source: McLean, 1988. Monitoring and Evaluation in the Management of Agricultural Research. Working Paper 14. ISNAR: The Hague

Table 2. Types of evaluations, in relation to object and level

Decision-making level	Object of evaluation							
	Context		Inputs		Processes		Products	
	Internal	External	Internal	External	Internal	External	Internal	External
System								
Institution								
Program								
Project								
Activity								

Table 3. Functions carried out by monitoring and evaluation at different management levels of the agricultural research system

Management levels	Types	Methods	Frequency (years)	Uses
1 Cabinet	Impact	Socioeconomic survey	10 to 15	Directs investment level toward broad areas
	Ex-ante	Technical and socioeconomic analysis	3 to 5	
2 Ministry of Agriculture	Impact	Socioeconomic survey	10 to 15	Guides the allocation of resources to research or development institutes
	Ex-ante Ex-post	Technical and socioeconomic analysis and reviews	3 to 5	
3 National agricultural research council	Ex-ante Ex-post	Technical and socioeconomic analysis and reviews	3 to 5	Determines potential impact of research initiatives to guide budget allocation for research institutions
4 Research institution	Ex-ante	Technical and socioeconomic analysis	3 to 5	Determines potential impact of research initiatives justifies/assigns resources to divisions/programs
	Monitoring	Evaluation of resources	Periodic	Raises the efficiency of management of the research institute
	On-going	Peer and expert review	Annual	Improves research implementation and planning
	Ex-post	Technical and socioeconomic review	3 to 5	Lessons learned: extension and the complementary character of programs
5 Research station	Monitoring	Research and procedure evaluation	Periodic	Improves the station's management
6 Program	Ex-ante	Technical and socioeconomic analysis	3 to 5	Determines potential impact of diverse approaches and research projects
	Monitoring	Resource and procedure evaluation	Periodic	Improves program management
	On-going	Peer and expert review	3 to 5	Guides short-term program planning
7 Project	Monitoring	Resource and procedure evaluation	Periodic	Improves program management
	On-going	Peer and expert review	Annual	Guides modification of on-going projects
	Ex-post	Technical and socioeconomic review	3 to 5	Guides toward future projects
8 Researcher	Monitoring	Research and procedure evaluation	Periodic	Improves activity management
	On-going	Peer and expert review	Annual	Guides research planning and execution

Source: McLean, 1988. Monitoring and Evaluation in the Management of Agricultural Research. Working Paper 14. ISNAR: The Hague.

Present Situation of Evaluation in Latin America and the Caribbean

In Latin America and the Caribbean (LAC), there is considerable experience in both research evaluation and institutional evaluation. The latter is richer in methods and mechanisms, and in the approaches and institutional structures put into practice by some institutions. Although there are many weaknesses and failures, there are also many positive experiences with evaluation in LAC. Appendix 2 presents a summary of the types of evaluation used by the agricultural research institutions in the region.

The following are the main points of this experience, as summarized by Novoa and Horton, 1994.

Experiences and methods

As in the case of planning, experiences with evaluation are closely related to the characteristics of the institutions, their mandates and fields of action. In the larger institutions, which address other fields besides research, and cover several products and regions, the experience in evaluation is richer and more varied, and the degree of progress in methods and procedures is greater. This is even more true when the institution has had previous systematic experience in planning. In some cases, progress has been made in distinguishing between institutional evaluation and agricultural research evaluation.

Even in cases such as the National Institute of Forestry, Agricultural, and Livestock Research (INIFAP), Mexico, where evaluation is said to not be institutionalized, the three main divisions—formed from previously existing entities which were merged—have defined evaluation mechanisms, as well as experience, especially at the program and project levels.

Peer review, internal and external review, technology validation at the farm level, and economic impact studies are the predominant types of evaluation in the cases studied.

The weak link

The abundance of evaluation studies can give the impression of a wide and generalized use of evaluation in the region. But this is not really true. **In nearly half of the organizations studied, evaluation is considered the weakest link in the general process of PM&E.** As an instrument of research management and administration, it is the least developed, being neither institutionalized nor well organized, nor differentiated from other components and actions.

The Sugar Industry Research Institute (SIRI, Jamaica), for example, clearly states that it has no formal type of research evaluation or external reviews, impact studies, or *ex-post* evaluations.

A similar situation occurs at the National Commission for Agricultural Research and Technology Transfer (CONITTA), Costa Rica. In this case, it is because the organization was formed only recently and essentially it co-ordinates and guides research conducted by other bodies.

At the Institute of Agricultural Science and Technology (ICTA, Guatemala), impact evaluation is only done at the farm level. Institutional research evaluation has consisted of external reviews. The institution's interest evaluation has focused on evaluating technology adoption by producers.

Similar situations are found at the National Coffee Research Center (CENICAFE, Colombia) and the Tropical Agriculture Research Center (CIAT, Bolivia). These institutions have a greater relative development and application of evaluation methods at the program and project levels. However, evaluation has been *ad-hoc*, conducted when special opportunities present themselves or external pressures or require it. Evaluation is not well related to institutional or research planning and monitoring.

In the recent past, the Colombian Institute of Agriculture and Livestock (ICA), has been outstanding for the large number of economic evaluations carried out. The institution has also had several global reviews organized by program and by discipline. As a result of these experiences, it has established a unit specializing in strategic planning and evaluation. However, the current situation is described as "deficient in any formal and permanent system of research evaluation."

Types of evaluation

External reviews are generally used by external groups to evaluate performance or ensure the good management of research. For this reason, donors frequently organize external reviews of the projects they finance. Sometimes external reviews are

organized for research institutions or national systems. This is generally connected with the identification of, or preparation for, technical assistance projects with external funding, for example by the World Bank, the Inter-American Development Bank (IDB), or the U.S. Agency for International Development (USAID).

In some large and decentralized institutions, such as the Brazilian Enterprise for Agricultural Research (EMBRAPA) and the National Institute for Agricultural Technology (INTA), Argentina, periodic reviews of the experiment stations and the research centers are organized to ensure the good management of these decentralized units. A cycle of external reviews is planned in these cases, whereby each unit is evaluated every 5 years. However, this cycle has not always been adhered to and some centers—in EMBRAPA, for example—have never had an external review. The external review of research programs is not common, unless there is external funding.

Internal reviews (generally annual) are common at the level of programs and research centers, but are rare at the level of institutions, except in the case of small institutions, such as CIAT-Bolivia. They are also rare at the project level.

Impact studies are not frequently done in the region; they originated mainly from the external demand for institutions and programs to validate results and justify resource allocation. Economic evaluations have generally focused on evaluating the costs and benefits of successful projects or programs. With the exception, perhaps, of EMBRAPA, they have not been institutionalized components of PM&E systems in agricultural research institutions.

Motivations and perspectives

In most LAC countries, there has been a growing participation by the private sector in agricultural research. Partly for this reason, public institutions are more inclined to take marketing conditions into account, both in respect of products and technologies. This has led to redefining research objectives and the criteria for evaluating the results.

Producers' associations and industry groups increasingly feel that they should be more involved in the whole process of technological development and modernization, from the formulation of policies and plans to the evaluation of results and benefits.

These signs of the times have been recognized by the institutions. Most of them are seeking to incorporate mechanisms in their operational schemes which will allow these sectors to participate more in their management processes and decisions.

This, together with the reviews of and changes made to the general PM&E process, has had the positive result, that evaluations tend to be more participative than in the past.

Case studies indicate that efforts to strengthen evaluation should be considered in the context of **decentralization**, **participation**, and the **use of results**. Currently, many research institutions are going through a process of decentralization; this implies the need to introduce periodic reviews or evaluations of the decentralized units.

Improved evaluation frameworks are needed, with clearer designs and terms of reference. These TOR should state the evaluation objectives, key questions, information needs, data sources, analysis methods and the intended audience and use of evaluation results.

Clearer terms of reference and designs are needed. These should state evaluation objectives, key questions, information needs, sources of data, analysis methods and intended use of evaluation results.

Evaluation should directly involve clients by having them participate in evaluating teams, or indirectly, by having evaluators to contact the users of research results. If evaluation is participatory, sharing results is part of the process. As a result, evaluation results will be better received, and used in decision making.

As was indicated in a meeting on the evaluation of agricultural research in LAC (Novoa, 1989), many evaluation results are not used because they are considered alien and foreign to immediate interests and to researchers' priorities. Others are not used simply because they never reach the beneficiary, and others because they are impossible to understand when they do reach them. Many of these problems originate in the evaluation methods used, in the way results are presented, and the elitist handling of information.

When this situation became evident in the Project Synthesis Workshop (Uribe and Horton, 1993), impact studies were considered to be more helpful if their results were explained in simpler, less academic terms, and if they reached a wider audience.

Examples in the region

As examples of impact evaluation experiences in LAC, the following cases are described in global terms: evaluation at the National Institute of Agricultural Research (INIA), Chile, evaluation at the producer level at ICTA, Guatemala, and economic evaluation at ICA, Colombia.

INIA, Chile: Economic evaluation. Some research programs or groups of activities at INIA have been the object of economic evaluation studies, using standardized procedures. The purpose of these studies was to determine the social and economic benefits generated by agricultural research.

The economic evaluations have been oriented towards determining the classic parameters in project evaluation, such as internal rate of return, and net present value. With this aim, streams of annual benefits from research are estimated (e.g. the value of increased production). Once the benefits of the use of new technologies has been determined, the costs associated with the generation and dissemination of these techniques are deduced, to arrive at the parameters previously designated.

Studies have been carried out to evaluate the generation of wheat, maize, barley, and rice varieties, as well as the introduction of bean seeds and biological pest control. Some of these studies have been carried out at the Catholic University of Chile and others at INIA as postgraduate studies.

Social evaluation studies at INIA have always been the result of external demands or scientific curiosity, and not of any intention to change institutional or political decisions. Thus, it is not surprising that they have been published as scientific papers, while other planning and monitoring documents are only circulated internally within the institution.

Many INIA researchers do not even know that such evaluation studies have been carried out. In the interviews for this case study, they were never referred to. For this reason, unfortunately, evaluations are not being used to increase staff commitment to the organization, nor to show the professionals at INIA the importance of their work.

Cost-benefit studies have been used by INIA to show the government why research is important. However, the studies have been criticized for being very general, for not having contact with agricultural researchers, and for lacking important information. These problems are inherent not is only in the Institute's studies, but also in its methods. Even though the economic evaluation studies are theoretically solid, are part of a strong tradition, and are associated with standardized procedures, they have not resulted in permanent internal management structures or systems, nor in organized teams of evaluation researchers at INIA.

Impact studies have also been carried out at INIA. The reduction of economic damage caused by an insect to Chilean fruit production was estimated before and after the use of a chemical control method recommended by the Institute. The increase in crop yields resulting from fertilizer use was also calculated. An attempt was made to relate this increase to the technologies developed by INIA, and to research and technology transfer activities carried out in Chile by other institutions.

The impact of specific technology transfer activities has also been evaluated. "Technology Transfer Groups" were evaluated in terms of their technical, economic, and social effects. The impact on maize and wheat production of national production contests, promoted by INIA and other institutions as a means of stimulating technology adoption and transfer among farmers, has also been evaluated.

Lastly, studies on the impact of demonstration centers created by INIA evaluated the benefits and identified the factors that could explain their success. (Borges-Andrade, 1993).

ICTA, Guatemala. Evaluation of results at the farm level. Part of ICTA's working methodology is to establish "trial plots" where the producer himself can evaluate new technology. Some time after establishing the plots (usually a year), an interdisciplinary group carries out an analysis called "acceptability evaluation," whose objectives are:

- To determine whether the technology is considered useful by a group of farmers.
- To classify the different practices according to their acceptability, and to determine the probability of their being widely adopted.

- To determine the reasons that collaborators had for adopting or rejecting a material or a practice.
- To provide feedback to researchers to confirm or reorient research.

These objectives are achieved by conducting surveys among farmers, in which the main variables refer to the number of farmers using the technology, the percentage of area under the technology, and the reasons for acceptance or rejection.

If the analyst considers that the new technology is used to an acceptable degree, the information is passed on to the extension agencies for dissemination. If is not acceptable, the technology-generation process returns to an earlier stage.

This method is part of ICTA's basic working strategy. It has been applied since the first technological recommendations began to come out, and has proved to be a useful feedback tool. The application and effectiveness depend greatly on the capacity of the Socioeconomic Unit, which has lost a lot of staff, especially at the regional level. When the evaluation is not carried out for lack of resources, the new adjustment of the technology has been observed to be negatively affected (Romano, 1994).

ICA, Colombia. Economic evaluation. ICA has carried out three types of economic evaluation of research and technology transfer at three levels: ad hoc, global, and dynamic.

A number of **ad hoc economic evaluations** were carried out at the beginning of the 1980s to evaluate the impact of the Green Revolution. The approach was to measure the effects derived from the increased yields of those crops for which improved varieties had been developed, distributed, and adopted.

In general terms, the evaluation analyzed the benefits for both producers and consumers. These benefits originated from a displacement of the supply curve, which reflected an increase in production and a decrease in market prices. This evaluation methodology saw research, transfer and adoption as part of an integral process, and gave the total effect of the benefits without discriminating between the various components.

Criticisms were heard outside ICA, and in some cases, within the institute's social science unit, of this methodology for excluding the social costs resulting from the intensive use of capital for applying the technological package accompanying the new varieties. It was also criticized for sectoral and regional imbalances, and because it tended to overestimate profitability, which had been very high in previous studies.

Ad hoc economic evaluations were carried out for cotton (1972); soy beans and rice (1973); wheat (1974); and palm, barley, and potato (1976). Other evaluations were done outside the Institute by doctoral students.

These ad hoc evaluations helped ICA demonstrate the returns to resources allocated to the core budget.

In his doctoral dissertation, Romano (1987) made a **global economic evaluation** of Colombia's agricultural research system. The approach was global and intersectorial. The evaluation's objectives were to (a) analyze the relation among different indicators of technological change in Colombian agriculture, for example, tendencies in productivity and use of inputs; (b) construct indexes of agricultural productivity based on the concept of total factor productivity; (c) estimate the average and marginal rates of return to public investment in the agricultural research system; (d) estimate the time log between initial research investment and its impact on production; and (e) estimate the indirect benefits of investment in agricultural research.

The evaluation methodology applied by Romano allowed him to arrive at three basic conclusions:

- The technological development of Colombian agriculture during the period 1960-1982 showed three phases: take-off during the sixties, dynamic growth during the seventies, and deceleration during the eighties. These phases were directly associated with total public investment in research.
- The growth of total factor productivity during this period was similar to that of the United States, i.e. 1.8% per annum.
- The technology developed, diffused, and adopted had a strong bias towards the intensive use of land.

In an extension to his doctoral dissertation, Romano (1987) developed a theoretical and empirical framework to estimate the indirect effects and the multipliers of technological change. In this work, he

evaluated the impact on non-agricultural production on family income and on employment.

As a component of an ICA-World Bank loan designed between 1984 and 1989, a **dynamic evaluation** system was designed to appraise the economic and social impacts of research and technology transfer activities.

The system's objectives were to evaluate mechanisms used in the technology transfer and the adoption of results at each stage of research, to detect the factors limiting those mechanisms, and to evaluate the economic impact of the technologies adopted. The system was to be incorporated as a permanent process, interrelated with the planning and monitoring processes.

The proposed system is made up of the following interrelated analytical components:

- Inventory of technologies generated.
- Economic analysis of technologies generated.
- Analysis of the socioeconomic environment.
- Estimation of the economic impact of the technologies adopted.
- Estimation of the adoption rate.
- Analysis of the generation, transfer, and adoption processes.

The proposed system is interesting because of its comprehensiveness. However, it was not implemented in ICA (Posada, 1994).

Summary

In this Sequence, evaluation was defined as a judgment, an appraisal, or an assessment of the merit, value, or quality of research, whether it is proposed, is on-going, or has been completed. Reference is made to the evaluation criteria of relevance, effectiveness, efficiency, and impact. Evaluation in an agricultural research institution can be conducted at several levels, from an experiment or work plan to the review of all the activities of an institution.

An analysis is made of how evaluation results can be used for different purposes, at different management levels, and at different research stages. The two most important uses are **public accountability and management decision making**. The first refers to an individual's or organization's responsibility to account for the appropriate use of resources. The second refers to the use of evaluation results in decisions making during planning, implementation, and periodic review of research activities.

Many objects can be evaluated, including: activities, projects, programs, centers or the institute itself. When defining an object for evaluation, its main purpose, assumptions, process, inputs, and products should be determined. It is also recommended that there should be a clear idea of the context in which the evaluation will be carried out.

Evaluations can be classified according to the moment in which they are done (**ex-ante**, **on-going**, or **ex-post)** and according to the point of view of the evaluator (**internal** or **external)**. Evaluation approaches included economic evaluation, **peer review** and **expert evaluation**.

Finally, the present situation of evaluation in Latin America and the Caribbean is analyzed. The use of evaluation and the methods employed are closely related to the characteristics of the research organizations, and their mandates and fields of action (e.g., private versus public; specialized versus broad mandate; and large versus small).

Sequence 2. Evaluation Methodology

Page

Flowchart for Sequence 2 24

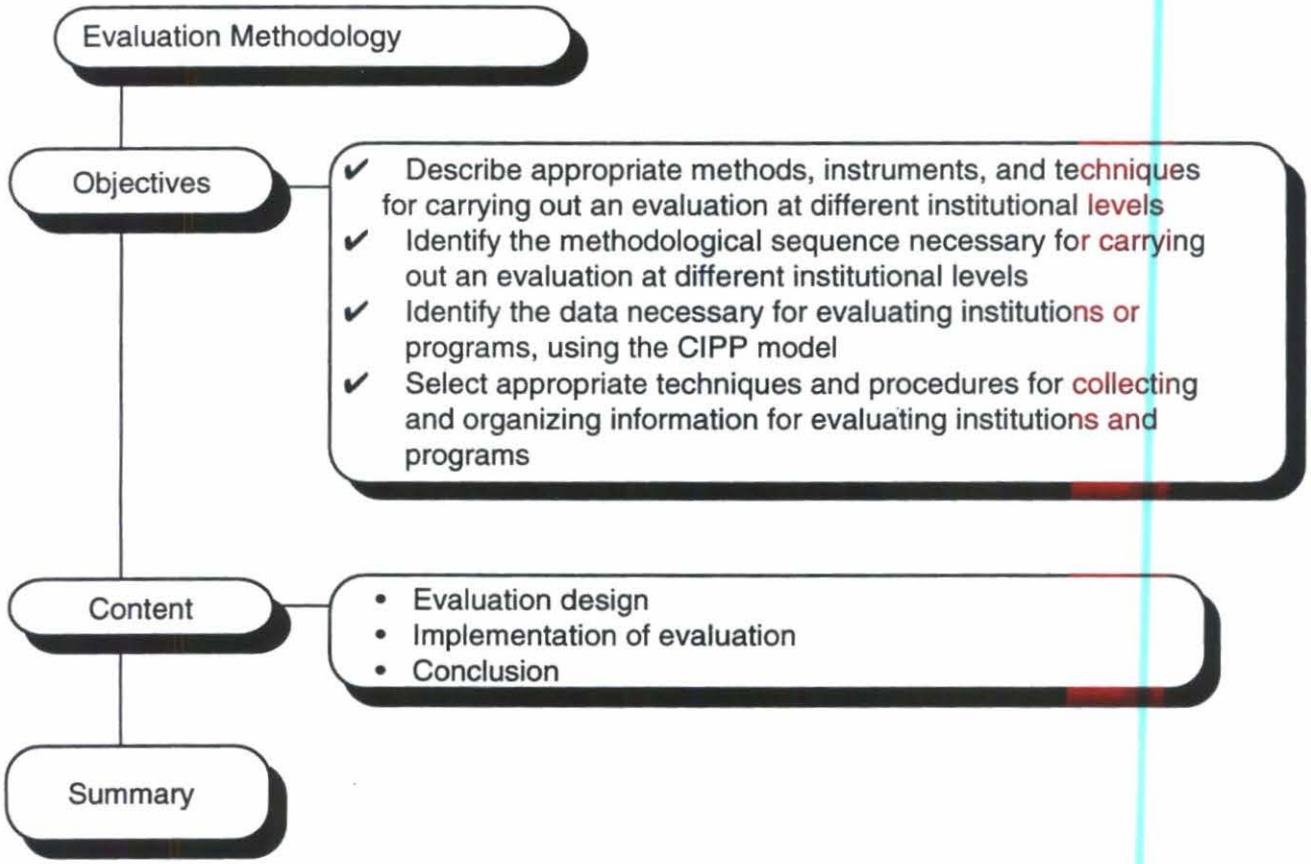
Evaluation Design 24
- Methodological sequences 25
- Research institutions as production systems 26
- Participatory identification of Context, Inputs, Processes, and Products 27
- Operational definition of variables 28

Implementation of Evaluation 28
- Instruments and procedures for collecting data 28
- Instruments and procedures for organizing data 32
- Information sources 33
- Data analysis 33
- Interpretation and critical analysis 36

Conclusion 37

Summary 37

Flowchart for Sequence 2



Evaluation Design

There is much to be gained if complex activities such as evaluations are planned in advance. Planning or designing an evaluation is an element in the success of all subsequent evaluation activities and it can help avoid wasting resources and effort. It can help guarantee that the data gathered are appropriate for the objectives of the evaluation. And it can allow evaluation results to be interpreted more logically and systematically.

Methodological sequences

One of the ways to plan and carry out an evaluation is to follow the methodology described below in five sequences. The sequences are applicable at different institutional programming levels. The rest of the sequence, however, emphasizes only those methodological aspects corresponding to the institutional and program levels, in the hope that these can be generalized for other levels.

1. **From objectives to concepts**
- Identify the objectives and/or priorities of the evaluation.
- Choose an appropriate evaluation model and the relevant concepts.
- Formulate causal hypotheses about their relationships, using the CIPP model as a guide.

2. **From concepts to operational definitions**
- Identify the variables that represent the concepts chosen.
- Define the variables operationally so they will be valid, believable, and feasible.

3. **From operational definitions to data collection**
- Define the sources for each variable, maximizing validity and minimizing costs.
- Choose the most appropriate instruments for collecting data at the highest measurement level and the lowest cost.

- Design instruments of measurements.
- Plan the data-collecting activity.
- Train data collectors.
- Manage the collecting activity, maximizing the validity, as well as time and cost-efficiency.
- Separate quantitative and qualitative data.
- Prepare the data analysis.
- Carry out analyses capable of revealing the meaning of the data.
- Prepare tables and graphs to present the data in an intelligible, succinct, and statistically correct form.
- Carry out the analyses to obtain the results needed for the original plan.

4. From data to interpretation

- Organize the qualitative data to describe the immediate context of evaluation, as well as aspects that can not be quantified appropriately.
- Analyze the descriptive data beginning with those relative to products and continuing with those referring to the other basic CIPP concepts.
- Analyze the causal relations among variables critically, using the causal hypothesis as a guide.
- Broaden the analysis by adding the unquantifiable aspects.
- Note interpretations, especially those which influence the more complex aspects of the situation.
- Make suggestions for making the most of opportunities and on the consequences of overall changes.
- Consider the results of the critical analysis from a strategic point of view, emphasizing important aspects, strong points and opportunities, without forgetting weak points and dangers.
- Summarize the written evaluation report.

5. From interpretation to presentation

- Finish with a specific presentation of findings and suggestions.
- Rigorously revise the form and the logic, avoiding confusing or unreliable steps.
- Write the executive report, omitting methodological details and emphasizing practical conclusions.
- Write a popular version, communicating the most important opportunities and changes found.
- Plan the distribution of the information and its use, making the message appropriate for each type of audience.

The main steps for carrying out an evaluation are:

- *Focus the evaluation*
- *Decide what will be measured*
- *Collect the information*
- *Interpret the evidence*
- *Write and present reports*

The previous steps can be applied with differing degrees of complexity and magnitude, depending on the size of the institution, program, or project, and on the resources available. However, the logic of the method must be followed, always answering the questions: what must we evaluate, why, how, when, with what, and what for?

Research institutions as production systems

In order to evaluate any organizational level such as an institution or a program or anything similar, we must take into account the fact that research institutions, just like the lower levels of the organization, can be understood as a production system whose product is technology, agricultural processes, or knowledge. This product is obtained through combining inputs in processes that take place within the organization to be analyzed and under the influence of an external environment (context) within which that organization operates (CIPP model).

Based on this conceptual model, the relevant aspects of the system can be selected for observation and analysis.

Concepts such as Context, Inputs, Processes, and Products are a useful simplification of the actual situation, but they are too far from the observation to be useful for the practice of evaluation. For this reason it is necessary to choose rather more concrete units of analysis to represent them. These units are called variables.

Selection of appropriate variables for an evaluation depends on techniques and creativity. The techniques are available in books on social sciences methodology. Creativity should be exercised by the evaluator to get the maximum benefit from evaluation situations to maximize the use of results, and to minimize costs.

Each of the four concepts (Context, Input, Processes, and Products of Evaluation) should be measured by a set of variables that are meaningful at the specific level being evaluated and which can be measured in all the possible conditions for collecting data.

A list of variables to represent the four concepts (Table 4) could include aspects such as:

Context. Potential users, interest groups, the program to which the project belongs, donor dependence, policies, and government priorities.

Inputs. Experimental designs the number and level of qualification of researchers, funds, quality of experimental fields, and the availability of information science backup.

Processes. Access to inputs, the level of researcher's liberty for decision making, the organizational structure of the unit, consultation styles among researchers, and communication processes.

Products. Research completed, scientific articles published, new technologies.

A presentation slightly different from that of Table 6 shows the relationship among variables, their operational definition, and procedures for collecting and organizing the information for the different institutional levels (Table 5).

The use of a basic model improves the understanding of the relationship among the parts being evaluated, so that the evaluation can be more than a measure of the levels of the variables, as will be seen shortly. The model takes the existence of products as an effect of the variables included in the other concepts. This difference in logical function among variables is represented by the use of the following classification:

- Dependent variables. Indicate the variables chosen as criteria for evaluation.
- Independent variables. Indicate the variables chosen as causes or important background.

These are functions that the variables take on in the logical context of analysis, because a variable is not intrinsically dependent or independent.

Table 4. Example of variables, levels, and operational definitions referring to different CIPP elements.

CIPP elements	Example of variables	Levels	Operational definitions
Products	New technologies Information	Program	Varieties and recommendations released during the period*
		Project	Advances in crossing varieties during the period
Processes	Quality control	Institution	Number of meetings of the quality committee during the period
		Project	Proportion of projects using appropriate scientific methods
Inputs	Quality of the human resources	Institution	Proportion of postgraduates among researchers
		Project	Appropriateness of the team's specialities for the project's needs
Context	Interest groups	Program	Main category of potential adopters (commercial producers or farmers)
		Institution	The most interested politicians in the region

* Percentage increase in the income level of adopters could be considered as an alternative.

Participatory identification of context, inputs, processes, and products

To evaluate is to judge the quality of the object in question. The concept of quality leads us to judge whether an object or event has been produced following previously established norms, either implicitly or explicitly defined. Determining adequate criteria for performing the evaluation is essential (Quirino and Coqueiro, 1985). Consultations with individuals and interest groups involved in some way in research are the basis for creating a participatory process of evaluation, and for identifying the crucial aspects to be examined. This may be achieved by submitting lists of variables to be reviewed and commented on by the groups, asking for their suggestions (Table 5).

The process of participatory evaluation with groups of producers is described below.

Participatory research with producers is a set at methods designed to allow their active contribution to decisions for planning and generating of agricultural technology. Evaluations with producers are a subset of these participatory methods.

Evaluation methods with producers can be used at different moments in the system outlined in Table 5: diagnosis, planning and design, experimentation, adaptation, and validation. These methods can be used in the diagnosis stage, to help producers express the criteria on which they base their decisions when faced with the alternative technologies presently available to them, such as choosing between different crops, varieties, or ploughing techniques.

Evaluation methods with producers can be used during the planning stage to preselect "prototype" technologies. This enables producers and scientists to decide together what technologies to validate.

Once the trials have been planned with producers, participatory research allows researchers and producers to generate and share systematic information on producers' reactions to the performance of the technology in the trials.

Table 5. Application of the evaluation with producers at different research stages

Research stages	Application
Diagnosis Identification of objectives, needs, and problems.	Identifies the producers' criteria for choosing available technologies, in order to understand his/her decision-making process
Planning and design Establishing priorities among problems, identifying potential solutions, designing "prototype" technologies and strategies to test them.	Identifies producers' reactions to prototypes, to set criteria for testing prototypes.
Experimentation Trials and evaluations of the "prototypes" transformed into developed technologies.	Identifies producers' criteria for choosing technology from among the options, with the purpose of selecting the most promising one(s).
Adaptation and validation The technology developed is tested and recommendations for its use are developed.	Verifies producers' reactions, comparing the new technology with current practices to ensure acceptable recommendations.

Source: Ashby, 1992

In the stage of validation and adaptation, evaluation with producers should continue, in order to verify the opinions and selection criteria obtained in the previous research stages. Evaluations with the producers in this final stage can be important for analyzing decisive criteria and technology characteristics which can only be identified quickly once the technology has been applied on a semi-commercial scale.

Operational definition of variables

Solid conclusions are impossible without translating theoretical concepts into observable variables. Therefore, the operational definition of variables should reflect the evaluation objectives, its design (model), and the data collection and analysis tools chosen.

A specific example is found in the following experience of EMBRAPA:

Example: Optional Definition of Variables at EMBRAPA

EMBRAPA carried out a global evaluation in 1991, in which the concepts of context, inputs, processes, and products were applied in order to guide the selection of variables, the collection of data, and the analysis of results. Five evaluation levels were used: projects, researchers, programs, units, and the organization as a whole. The analysis of research projects and programs included a list of variables that served as indicators for each of the four concepts:

Performance indicators (Products). Examples: (a) length of project; (b) research progress versus goals; (c) fulfillment of deadlines for completion, and disbursement of funds; and (d) expected outputs.

Input variables. Examples: (a) the qualifications and number of researchers; (b) qualifications and number of support staff; (c) qualifications of physical and resources space; (d) funding availability and sources; (e) technological factors used in research; and (f) sources of scientific information.

Organizational variables (Processes). Examples: (a) reporting relationships; (b) hierarchical levels of decision making; (c) relationships between researchers and supervisors; (d) technical-scientific and management qualifications of directors; (e) internal and external communications systems; (f) internal and external information systems; (g) technical-scientific and administrative information system; (h) the unit's management style; (i) organizational and functional systems; and systems for assigning responsibilities and levels of autonomy; and (j) organization of human resources by teams.

Variables of the external environment (Context). Examples: (a) potential users and fulfillment of society's needs; (b) national research program to which the project belongs; (c) characteristics of the national programming system; (d) region in which the project is inscribed; (e) dependence on national and international donor institutions, and on scientific and technological interchange; (f) relation of the project to EMBRAPA's policies, to agricultural, scientific, and technological policies, and to budget guidelines; and (g) technology transfer provided (EMBRAPA/SEA, 1990).

Appendix 3 presents several classifications, for variables.

Implementation of Evaluation

The implementation of an evaluation requires care in the selection of instruments and procedures for collecting data.

Instruments and procedures for collecting data

First, the operational definitions of the variables must be transformed into collection instruments, such as questionnaires, summary sheets, interview schedules, and other guides. Collection procedures depend on those who provide the information and, therefore, on where they are situated within the organization. Adjustments need to be made so that the instruments are adapted to the characteristics and needs of those who use them.

In every case, data collection methods and instruments should guarantee the following qualities:

Evaluation methods are **valid** if they measure what they claim to measure. Since decisions affecting people's lives and jobs may be based on the results of an evaluation, it is very important that the methods be reliable. To ensure the validity of results

it is recommended that more than one method be used in the evaluation process.

Evaluation methods are **credible** if the people for whom the evaluation is done accept them and believe in their usefulness. An evaluation's results are not likely to be applied if the decision-makers do not understand and trust the methods. For this reason, administrators often ignore the results of evaluations using sophisticated, but confusing, methodologies.

Finally, evaluation methods are **feasible** if they can be implemented in the specific organizational context proposed for their use. Some reliable, credible methods are too expensive, take too much time, or are too complex to be used in practice (Horton *et al.*, 1993).

- *Methods are **valid** if they measure what they claim to measure.*
- *Methods are **credible** if people accept them and believe in them*
- *Methods are **feasible** if they can be implemented*

Different measurement levels can be distinguished. The most basic one distinguishes observations by quality but not by quantity. This level of measurement gives nominal scales which place objects into mutually exclusive categories.

Quantitative measurement gives scales that distinguish categories by order (ordinal scales), places them at fixed distances (interval scales), or uses a fixed point at the beginning of the scale that can logically be identified as zero (rational scales).

Measurement levels limit the use of statistical treatment alternatives appropriate for the data to a certain extent, but creative statistical techniques accept, for example, treatment of nominal data in statistical trials designed for higher levels of measurement. This is done using binary variables in which 1 (one) represents the presence of an attribute and 0 (zero) its absence (Table 6).

Qualitative measurement is necessary, convenient, and advisable for many variables and many opportunities, but validity, credibility, and feasibility should not be sacrified. With care, it is possible to make good evaluations using mainly qualitative data. On-going evaluations, for example, tend to use mainly qualitative data.

Quantitative and qualitative variables

Analysis of quantitative variables is mainly done using statistical means. Qualitative variables need not be excluded from the analysis, because they can be treated as binary or classifying variables. They can be included in tables, graphs, and even in more complex and powerful statistical treatments.

Quantitative data are often used to describe the context of the units being evaluated. For example, a program evaluation needs a description of the organizational context in which it is established. Evaluation of a research system (NARI) needs data on the country, its government, and its agriculture. For this, variables are used that were collected and interpreted at levels different from the one being evaluated. Many of them will only be qualitative variables.

Instruments for collecting data should make the most of the information, in order to reach the highest levels of measurement. Higher level scales

Table 6. Levels of variable measurement for evaluation

Levels	Variables	Operational definitions
Nominal	The Center's Production	New cultivars were released last year
Ordinal	The Center's Production	Classifying the Centers by the number of new cultivars released last year
Intervals	Temperature	Monthly averages of maximum and minimum temperature in degrees centigrade
Rational	The Center's Production	Number of cultivars released last year by each Center

of measurement can be transformed into lower-level scales but the converse is rarely true.

Data are usually classified according to their availability for the research to be done, which in this case refers to evaluation (a form of social science research).

Data collection instruments are means for registering information conveniently to attain the highest measurement levels. Scales of the highest measurement levels can be transformed into lower-level scales, but the converse is rarely true.

Primary data are those collected directly for the evaluation. Collection instruments to be used can be designed from the outset, determining the characteristics recommended, such as the level of the variables. But it is not always necessary to collect information from the beginning, because it is possible to use secondary data.

Secondary data are those taken from other sources previously available. Thus, their desirable measurement characteristics are beyond the control of the evaluators. Secondary data for an evaluation may be data collected during monitoring. The compensation is that they are generally cheaper and less effort is required to get them.

Primary data

Primary data collection begins with the operational definition of the variable. Generally, these data are collected directly at the organizational level to be evaluated. Sometimes data needs to be collected from smaller units and aggregated or consolidated to the level of the unit being analyzed. In the design of instruments for collecting primary data, the characteristics of the variables specified in Appendix 3 are taken into account. In general, the sources of information are people, but primary data can also be obtained from observations, text analysis, or another form of collecting information that proceeds directly from the actual situation.

The ideal is to limit to a minimum the design of instruments for collecting primary data, and to use data from monitoring, which if well-planned and co-ordinated, can meet a large part of the data needs of an evaluation.

Secondary data

Secondary data are very important for evaluation. They represent the set of knowledge collected on previous occasions for other ends. The records coming from systematic monitoring are probably the most valuable secondary data for evaluation, along with information coming from the planning process. Other sources of data, such as institutional records and documents, previous evaluation reports, and studies files, provide abundant secondary data.

In economic evaluation, secondary data are often extensively used. However, it is generally important to complement them with primary data too.

Designing instruments for collecting data

Designing instruments for collecting data is both a science and an art. Social science libraries have specialized literature that can be very useful (Duverger, 1981; Pardinas, 1977; Martínez *et al*, 1983). It is important to remember that improving the quality of the instruments, can improve the quality of the data, and therefore, of the evaluation.

There are three main types of instruments for collecting data:

Observation guides are used to observe behaviors or situations without the interference of informers. They are useful for observing cases, as, for example, field experiments, soil conservation, greenhouses, etc., and the ability of lab technicians. These should include a list of the objects and/or behaviors to be observed, and a space for recording measurements. A checklist can be used for recording variables that require measuring on a nominal scale, or with a scale of intensity for measurements of greater precision (Table 7). Observation guides are very useful for organizing work in external evaluations.

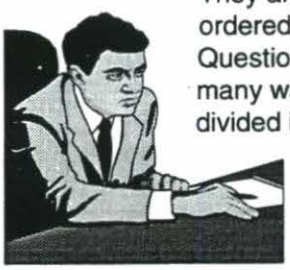
Observation of behavior or situations without needing informants

Interviews are used to collect very specific (not repeatable) or sensitive information that must be clarified and detailed at the moment of collection.

They are useful for data based on the personal judgment of privileged observers, such as bosses, previous directors, authorities, and frequent participants in strategic events. The interview should be prepared in advance to guarantee that all relevant aspects are considered; it can be structured as a written script, leaving spaces for recording the information.

Collecting very specific or personal information

Questionnaires are used for collecting the same information from many people; they are useful for facilitating answers and their statistical treatment. They are a collection of questions ordered in a logical sequence. Questions can be formulated in many ways, but generally they are divided into open and closed questions. Open questions do not specify alternatives for answering. This makes interpretation difficult due to the lack of precision in the information given. Closed questions offer categories for answering that must be marked by the person being questioned. These are difficult to construct and, if they do not include the necessary alternatives for all variations possible among those surveyed, the answers may give a wrong impression. However, results obtained are easier to compare and interpret statistically. They are frequently used in internal evaluations. The pretest for this Module has examples of open and closed questions.



Collecting data from many people

Procedures for collecting data

Access to the people and the type of instrument chosen determines the way the data will be collected.

Observations are always carried out by people who have access to the environment to be evaluated. When different observers are working separately, interpretation of the variables to be measured and their respective operational definitions should be standardized beforehand.

Personal data are collected by means of interviews or questionnaires. The relationship between the interviewer and person interviewed is very important for obtaining reliable information (see Appendix 4). Interviews can be carried out over the telephone; this reduces travelling costs to far-away or not easily accessible places. The mail and other similar means of communication are also useful and economical, but the time needed for receiving the reply must be taken into account.

Group interviews are very useful for discussing aspects related to the motivation for and judgments of complex processes that cannot be totally visualized, or for when seeking the collective, formal participation of interested groups (researchers, users, opinion groups, etc.). Electronic means is used more extensively every day. Computers and fax machines are helpful in collecting data; the latter is a substitute for conventional mail for transmitting information.

Computers offer ample opportunities for data processing and for increasing precision. For example, in the evaluation of research projects carried out by EMBRAPA (Quirino, *et al*, 1992), data were collected electronically in December 1990. After defining the questionnaire, using operational definitions of the chosen variables as the basis, an interactive program was established for collecting data. Those interviewed were high-level researchers (associate center directors, program coordinators, and administrative researchers) who had access to a computer. They received disks with the program, which they later returned with the answers. They read the questions on screen and typed the answers within the acceptable alternatives according to the range of the answers.

Information was then duplicated and handled electronically from the disks, thus avoiding errors caused by excessive manipulation and data transmission between media.

Psychological aspects of collecting data

Every time that data are collected for research, validity, credibility, and feasibility must be taken into account. This also applies to evaluation data. The evaluator must handle the relationship with the person interviewed in such a way that it allows to get the best quality information possible. Collecting data represents an opportunity to create

Table 7. Observation guide for the laboratory (sample)

Activity: *Estimating stomach contents of mites natural enemies*

Institution: *CIAT*

Program: *Importance in cassava of mites' natural enemies*

Unit: *Acarology*

Observer: _____

Observed: *Preparation and gel run for electrophoresis* Date:

Components/tasks	YES	NO	Scale 0	1	2	3
1 Were testing units, carefully prepared?						
2 Were concentrations of reactives for preparing the gel well determined?						
3 Were amounts of reactives required measured correctly?						
4 Was the gel poured correctly?						
5 Was the lab equipment handled carefully?						
6 Were mite samples sufficiently macerated?						
7 Was time of gel run correctly estimated?						
8 Was the gel stain well prepared?						

Observations _____

Scale: 0 = bad, inadequate
1 = poor, deficient
2 = good, acceptable
3 = very good, highly satisfactory

expectations about the institution and the planning, monitoring, and evaluation process. Many opportunities arise for creating incentives and increasing participation in PM&E among internal and external groups and individuals. Groups and persons interviewed should be sent feedback of information available from results of the evaluation. These are the formative aspects of data collection.

Instruments and procedures for organizing data

In general, more data are collected for evaluation than are really used. This could be avoided by focusing data collection on key questions and on the context, input, process, and product variables discussed previously. Data do not always refer specifically to the unit or level being evaluated. For example, to evaluate the quality of a program, data are also collected on national agricultural research budget to analyze the context in which the program operates.

After identifying the organizational level to which the data refer, these should be organized in a grid of columns and rows. The rows identify the units of reference, and the columns identify the variables they measure (Table 8).

For data to be useful, their role in the evaluation logic must be determined.

This form of organizing data logically, dividing them by level and by variable, helps the evaluator distinguish between information that is only useful for describing the units evaluated within more encompassing levels of the organizational environment, and the information that describes and compares units that are the central theme of the evaluation.

Information of the variable N related to the program X is located in the intersecting cell (XN). This leads the evaluator to think in terms of the level of data abstraction and of their relevance, thus avoiding collecting data that will not be used. This guarantees adequate information processing.

Table 8. Grid to guide the organization of data for a program evaluation

	Variables										
Programs	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10...**	**N**
Wheat											
Rice											
Maize											
Beef cattle											
Flowers											
.											
.											
.											
.											
.											
Xn											

On the other hand, a slightly different presentation from that in Table 4 emphasizes the relation among variables, their definition, and procedures for collecting and organizing information at different institutional program levels (Table 9).

Electronic data processing is recommended, and is generally feasible; however, they can also be classified manually when they are not complex and numerous (Appendix 5).

Information sources

Information sources are chosen for their potential as suppliers of the information needed—whether because they are familiar with the aspect to be measured, or because they possess the relevant data—(based on the operational definition of the variables.)

The degree of interest and the availability of data will vary among sources. Internal informants who belong to the organization and are found at different management levels among researchers and support staff, will feel more obliged to co-operate, but may have an interest in biasing the information. If in doubt, design check of the variables by collecting data from two or more different sources. External informants may be users, beneficiaries, scientists, donors, or legislators. They might be less willing to inform, since they will not feel obliged to co-operate with the evaluation. The evaluator should keep these factors in mind and try to prepare informants psychologically so they will provide the best possible internal and external information. The data collecting instrument should also be adapted to the different informants, takin into account their capacity to understand, their willingness, and their ability to provide information.

Data analysis

No matter how data are treated, they must be synthesized to be interpreted and communicated easily. The measurement level applied limits the alternatives for systematizing the data and statistical

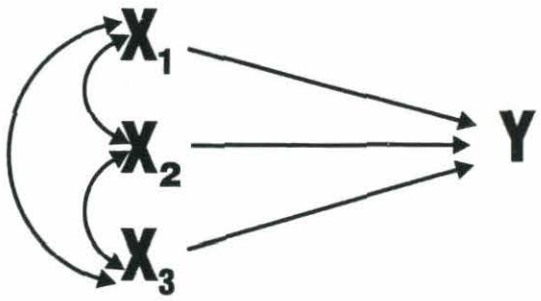
Table 9. Examples of variables, operational definition, and procedures for collecting and organizing data according to levels and CIPP elements

Levels	CIPP Elements: Variable		Operational definition of variables	Procedure for collecting and organizing data
System	Input:	Financial resources for research in the country	Amount of money assigned to the national budget	Analysis of national and institutional budgets; formation of a historical data base
Institute	Process:	Quality control	Number of meetings of the quality control committee	Analysis of minutes of committee meetings
	Input:	Quality of human resources	Proportion of postgraduates among researchers	Records of training unit, organization of a data base file
	Context:	Group interested in research	Most interested politicians in the region	Interviews, correspondence; organization of a directory
Program	Product:	New technologies	Number of varieties and recommendations released	Questionnaire for researchers, records of seed unit
	Context:	Interest group	Potential adopters: small farmers, commercial farmers organization of dynamic file	Interviews to extension agents and technical assistants,
Project	Product:	New technologies	Progress in varietal crossing	Interviews with researchers
	Process:	Quality control	Number of projects that follow the scientific method	Visits, review of project protocols

treatment. Frequency distributions are appropriate for nominal data, while tables of means and standard deviations appropriate for present data taken with more precise measurement. In both cases it is possible to differentiate between programs or categories using statistical tests common in social research manuals, such as chi squared (X^2), the difference of means, or Pearson's correlation.

It is easier to understand and communicate information if the data are presented in graphs with different formats, such as sector or bar graphs. Computer programs can help produce these graphs.

Relation between data and the concepts of context, input, process, and product. In data processing, two phases should be distinguished. The first describes the status of what is being evaluated, using, for instance, the CIPP concepts. The variables that correspond to each of these concepts are examined in sequence to construct a coherent image of the set.
The second phase of data processing is analytical, and is based on the relationships between variables, especially those in which the product variables are analyzed in relation to input, process, and context variables. This phase can begin by analyzing pairs of variables; one of them refers to one of the three antecedent concepts, and the other refers to the product concept. As a final synthesis, more sophisticated statistical techniques can be applied, using as the basis a causal model such as the following:



where X_3 = context
X_1 = input
X_2 = process
Y = product

Straight arrows indicate causality and curves indicate correlation. Figure 7 complements this.

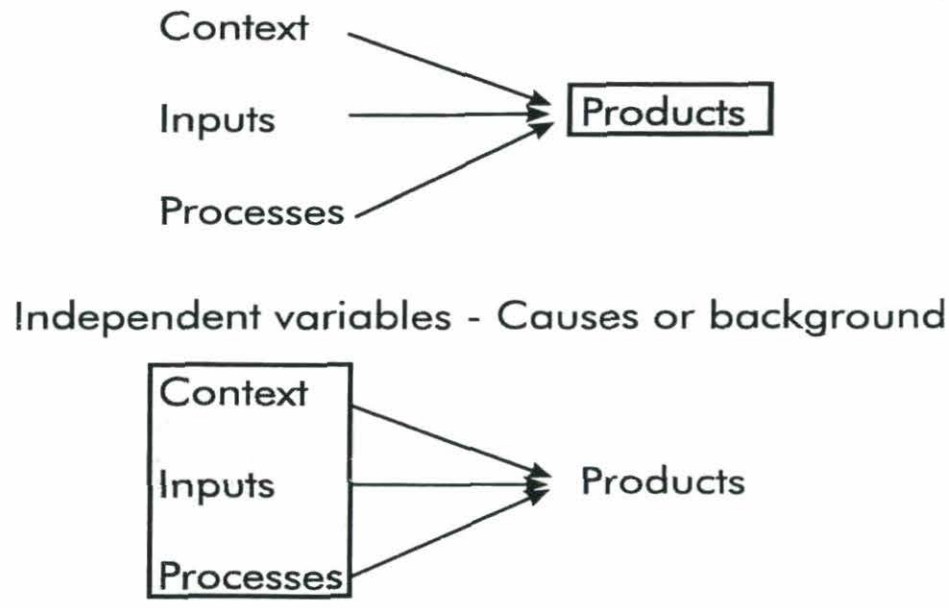


Figure 7. Relationships among variables

Alternative models can be constructed and different explanatory schemes can be tried to obtain evidence for prescribing new and more efficient combinations of factors in the institution. The order used in the data analysis can be modified for presenting the final report, but the data should always be explored to its broadest possibilities, within the limits of time and resources.

In a project evaluation, EMBRAPA applied the following sequence of data analysis (Quirino, *et al.*, 1992), beginning with the punctual and descriptive aspects and going on to the causal relations. Each project was examined, keeping in mind its relation to each aspect analyzed.

Research projects in the organizational context

- Planning phase
  - Needs assessment.
  - Clarity and conciseness.
  - Suitability of research procedures in relation to objectives.
  - Social impact.
  - Quality of planning.

- Execution phase.
  - Access to inputs.
  - Co-ordination among institutions.
  - Project alterations.
  - Demand for resources.
  - Actions to disseminate results.

Research projects in the context of society

- Access to resources.
- Users and products.
- Regional differences, for example:
  - Regional suitability of the projects.
  - Regions and different users.
  - Resources and regions.
  - Regional base of research products.
  - Regional differences in diffusion of results.

The performance problem

- Performance factors.
- Influence of internal and external validity and of formal political-institutional suitability on projected and achieved performance.
- Variation of the performance factors.
- Influence of inputs, processes, and external environment on performance.

Interpretation and critical analysis

The results of data processing are inputs for interpretation and critical analysis; and this is the most important stage of evaluation.

Interpretation and critical analysis seek answers to questions that were defined as priorities when the evaluation was planned. Results from the data should be carefully considered in order to grasp their meaning, and to deduce the practical consequences for the situations on which the evaluation was based. Practice shows it is easier to begin interpretation and critical analysis with specific aspects, but it is necessary to relate them with one another in order to find ideas for changes that would help solve any problems that may have been detected.

The last logical step of the evaluation is to present a synthesis of the problems detected and to propose technical administrative, political, economic, or training measures to solve them. In this step, the evaluators' experience and creativity play an important role, justifying the efforts made to hire them.

Often this stage of the evaluation does not receive adequate attention, reducing the validity and usefulness of the evaluation results.

If the evaluators are trained and have the means to do so, they can use analytical techniques such as path analysis (Briones, 1982), multiple regression and coefficient correlation analysis, looking for the most important influences that independent variables exert on dependent variables (Blalock, 1968).

These procedures are even more important in impact or economic evaluations in which the contribution of what is being evaluated and other social and economic factors needs to be identified. Lack of clarity hinders the practical application of the studies' conclusions.

In consequence, impact analysis is especially sensitive to problems of poor identification of the model, due to the multiplicity of variables— frequently unknown— that influence the social impact of programs and projects.

There are many analytical constraints to impact evaluation. First, the social world is complex and most social phenomena have many roots and causes. With so many factors involved, the severity of a social problem can be influenced by a greater number of causes than those modified by a program.

Secondly, models suitable for impact evaluation are difficult to develop because social science theories and empirical generalizations are still relatively weak and incomplete.

Third, only modest impacts are normally expected from social programs. The effects of programs are usually minor and therefore difficult to detect. Also, some social programs are especially difficult to evaluate because they have been in operation for a long time. On-going programs, covering vast populations, can only be evaluated if based on theoretical assumptions that often reduce credibility (Rossi and Freeman, 1985).

Conclusion

Developing an appropriate evaluation methodology requires the resolution of issues at three levels:

- At the political-administrative level, one must take into account the relations between evaluation and external groups, such as the government, donors, public opinion, interest groups, and society.
- At the administrative-organizational level, one must consider mobilizing internal resources of the organization and making them available to, and co-operative with, the evaluators.
- At the technical-scientific level, one must carry out the evaluation correctly from the point of view of social science techniques.

The last level is generally the weakest one. The risks are enormous. On one hand, evaluation can produce unreliable results that could harm some people unjustly. On the other hand, it can degenerate into a simple periodic ritual in which the perception of reality is manipulated to the advantage of some, with adverse effects on society and equality.

Top management levels must take care that all requirements of the technical-scientific level are met.

The interaction required between agricultural research evaluation and social research techniques is just beginning. This field needs the dedication of experts, the leadership of institutions, and the attention of donors.

Summary

Design is an essential precondition for guaranteeing the success of all the following actions in an evaluation. Agricultural research institutions can be viewed as production systems whose products are technology, agricultural processes, or knowledge. These products are obtained through the combination of inputs and organizational processes in the context of an external environment. Evaluation design should follow a methodological sequence which begins by identifying the objects to be evaluated and the reasons for the evaluation, taking into consideration the uses foreseen for results. Individuals and interest groups should be consulted, to create a participatory process.

Evaluation design involves choosing variables that measure concepts. Variables should be defined operationally, taking into account the organizational context in which they will be measured.

Carrying out an evaluation involves different stages: the first is to focus on the key evaluation questions; the second is to decide what is to be measured, to select appropriate instruments for collecting data, to ensure their validity, credibility, and feasibility in the evaluation process. In selecting the instruments for data collection, one should take into account how they will be used: for collecting primary or secondary data; if informants are internal or external to the organization; the level of precision sought (quantitative or qualitative); and collection procedures used (e.g. personal, by mail).

Data are collected in the third stage and organized for later statistical and logical analysis.

In the fourth stage, the simpler statistical results are analyzed to describe the status of the organization or process being evaluated, and to compare it with the ideal to be achieved.

Finally, evaluation reports are written and presented, taking into account the audience they are intended for and the information they need.

Sequence 3. Relating to other Institutional Processes

Page

Flowchart for Sequence 3 40

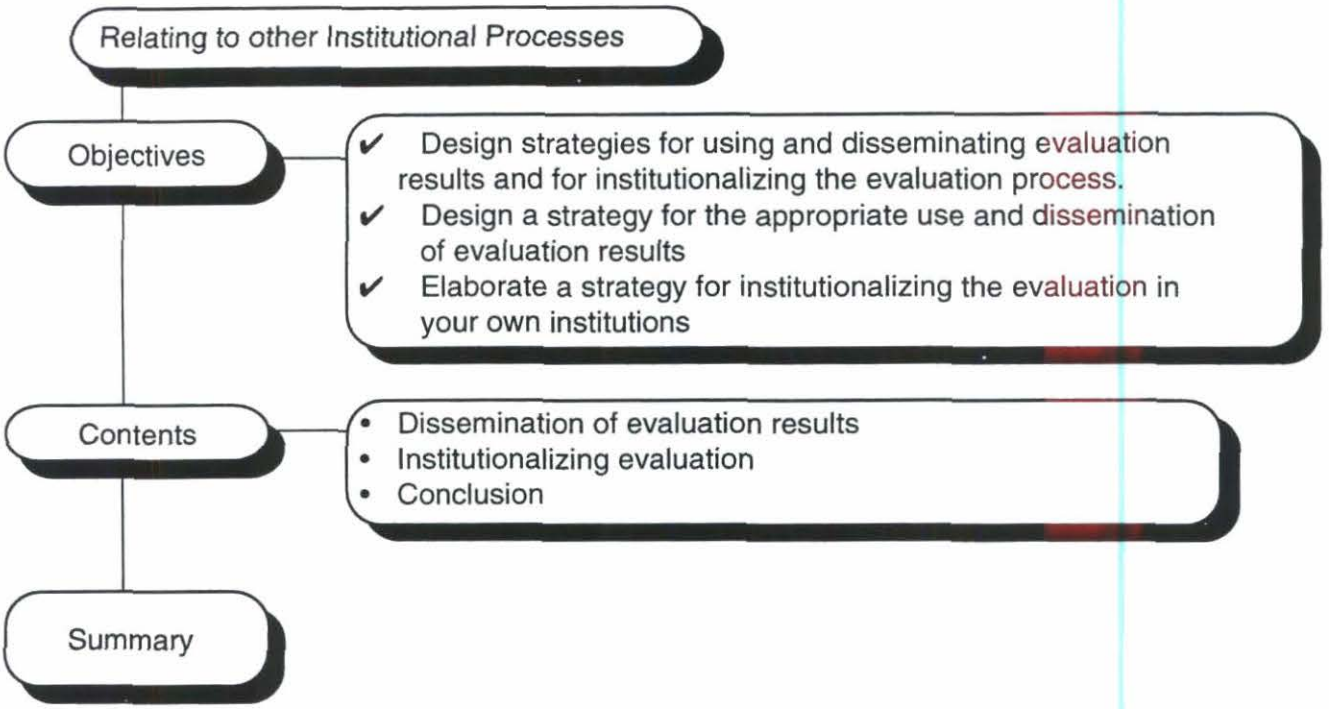
Dissemination of evaluation results 40
- Reports 40
- Uses of evaluation results 41
- Meta-evaluation 44

Institutionalizing evaluation 44
- Periodicity 44
- Institutional structure 45
- Evaluation capacity 46

Conclusion 49

Summary 49

Flowchart for Sequence 3

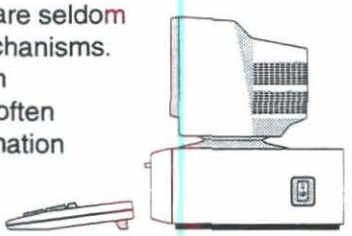


Dissemination of Evaluation Results

In an agricultural research organization, information should flow both "vertically" and "horizontally." On the horizontal plane, planning results (clear objectives, milestones, and resource allocations) are necessary for guiding the execution of research, as well as for monitoring and evaluation. Evaluation results, in turn can be useful for planning future research.

On the vertical plane, clear guidelines from the institutional level should guide middle-management's decisions. Then, information on specific research activities should flow upwards to enrich decisions at the higher levels.

Comprehensive plans (macro level) should be broken up into more specific objectives and work plans (micro level). On the other hand, plans and results of individual projects need to be synthesized, condensed, and translated into social-economic terms so that they can be used by institution directors, ministry officers, and interest groups.



Reports and meetings are seldom seen as evaluation mechanisms. However, with minimum restructuring, they can often produce valuable information that can be used for evaluation purposes. Perfecting existing mechanisms is one way to strengthen evaluation at a low cost (Fonseca, 1990).

Reports

Reports are often taken for granted. Little thought or effort goes into their design or into the preparation of information dissemination strategies for evaluation results. As a result, the quality and effectiveness of communications between evaluators, decision makers and those being evaluated is often poor.

Clear, precise, suitable communication

Different groups need different information in different types of report, As a rule, one general report prepared for several different groups, organizations, or audiences is not very effective. Instead, "specific reports" should be prepared for each audience, structured according to their interests and information needs.

One of the most common disappointments after an evaluation has been completed and once the report has been turned in, is the lack of attention to follow-up on recommendations, and the lack of action taken.

Evaluators can encourage the use of results by proposing alternatives for action and mechanisms to monitor the use of evaluation results.

Implementation of evaluation recommendations depends, to a great extent, on the consolidation of an institutional evaluation system. For this reason, when those at higher levels of responsibility receive the recommendations, they should work together to analyze them and propose plans of action.

Reports should be designed with the following key points in mind:

- Who should receive the report?
- What type of information should the report contain?
- How should the report be presented (written, verbal)?
- How can different stakeholders be helped to interpret and use the report?
- When should the report occur (beginning, middle or end of the evaluation process)?

Three key aspects should be considered in the process of communicating evaluation results: the audience, the type of presentation, and its contents.

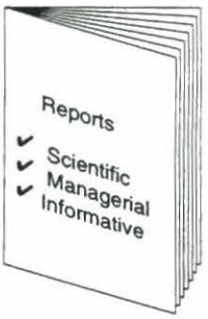
The **audience**, (for whom the information was prepared) will depend on what is being evaluated (as this will condition the information offered) and the purpose of the evaluation. Concerning **presentation**, there are three types of reports: scientific or formal reports, executive reports, and public awareness reports for distribution. **Report contents**, refer to the amount of scientific and methodological information and language used. This should depend on the kind of report being presented.

Scientific reports

A formal or scientific report is generally a reference document that explains in detail what was done and how it was done . This report discusses the methodology used, annexes information collection instruments, and documents the analysis in detail. It is useful as a reference for future references. As such, it should include a table of contents, summary, introduction, methodology, results, conclusions and recommendations, bibliographic references, and appendices.

Executive reports

An executive report can be a stand-alone document, or it can be the summary in a formal report. The executive report is more useful for a wider, but qualified, audience; it contains a clear and concise summary of the evaluation purpose, analysis, conclusions and recommendations.



Public awareness report

This report is directed to a wider public, and it contains only the most essential findings of the evaluation. It does not address technical, methodological issues, and is written in a simple, everyday language.

Uses of evaluation results

The evaluation can be directed towards diverse users, as for example producers' associations, government entities, funding groups outside the organization or research system, management at different levels of the organization or system, external development or funding agencies, or the researchers themselves. The type of report prepared should depend on the intended user.

When designing an evaluation, it is useful to identify and categorize possible users of results, and determine the information that will be needed. Detailed examples of the main internal users follow:

For leaders and policy-makers, involved in research, before establishing a research plan:

Results will be used to decide program design. A study of the complete situation of the research system in a country will be necessary in the earliest phase of developing of a global research plan. The same will be true if the

government decides to revise its development priorities. Such a study will also be necessary if there is a drastic change in research capacity, in technological potential, or in any other factor that modifies agricultural potential, and therefore research requirements. Government authorities apply results of this type of study to decide on research priorities, in the context of development objectives.

For research leaders, to select objectives and programs:

Results will be used to decide which programs will be designed, or whether it is convenient to continue with a proposed program.

In this stage, evaluation activities include more detailed analyses of problems, selection of research approaches, and review of the necessary and available resources (personal, infrastructure, budget, etc.)

For program coordinators, during the execution of a research activity (work plan, project, program, etc.):

Evaluation results are used to check that execution proceeds according to plan; to discover potential problems; and to adjust programming according to results. The users of this information are those involved in executing the evaluation, unless the problem discovered is too great for their management capacity.

For research directors, when an activity is completed:

Evaluation results are needed to appraise the results achieved and to understand the factors that influenced them; to assure that results remain available for future use to all those involved in research planning.

For leaders and policy-makers after a research activity is over: *Evaluation results are needed to estimate the contribution of research to development; to understand which factors (internal and external) influenced adoption results and impact; and to extract lessons for future planning. Users of this information are top-level research directors and those in charge of national policies. (Murphy, 1993).*

Currently, evaluation systems designed for internal management needs are less developed than those which respond to the needs of external funding sources (domestic and foreign). However, some NARIs which do not apply formal research evaluation systems have organized internal evaluations. Such is the case of INIFAP - Mexico. Its internal evaluations provide information on inputs used, institutional capacity at the program level, and results. CENICAFE-Colombia carries out annual evaluations to analyze the fulfillment of the work plans and programs. CIAT-Bolivia has carried out internal reviews when problems were detected, performance evaluations of its staff, and on-farm evaluations.

It is very important to determine for whom the evaluation is being done (Figure 8). The most important sources of information on expectations are discussions held with the stakeholders in the organization, and an analysis of the evaluation's purpose.

Oral presentations on the evaluation are effective to inform interest groups of results and to discuss the evaluation methods and conclusions. It is important to anticipate the questions stakeholders will ask about the evaluation, and to try to answer them. The evaluation report can only have an impact on decision making if it addresses the most important questions.

When done properly, an evaluation will provide useful information that can improve the management of programs, centers, etc., thereby increasing the possibilities of fulfilling social goals. Evaluation results often transcend the scope of the "unit being evaluated," establishing themselves as the basis of institutional planning, organization, and management. They can also serve as an input for the definition of regional, sectoral, and national policies. External evaluations carried out in Costa Rica by ISNAR are an illustration of the latter (1981-1986).

Unfortunately, diffusion of evaluation results is often restricted to direct users, reaching neither the policymakers nor external groups. As a result, opportunities to consolidate strategies and agrarian policies based on research experiences and on contact with producers, and to unify public opinion around research needs and achievements are lost. (Novoa, 1989).

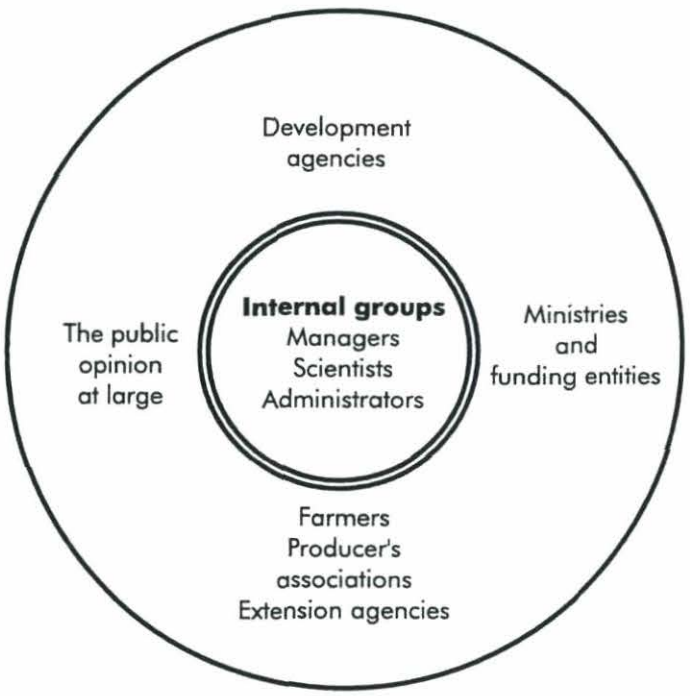


Figure 8. Who is the evaluation for?

Results from evaluation of agricultural research projects should have a broad dissemination. Policymakers need to learn about the benefits and impact of agricultural research. This can enhance the system's funding, stability, and its ability to face the challenges of agricultural development, a basic component in improving the welfare of the rural populations and in promoting economic development.

Information for decision-making

The need for external accountability is closely linked to the need for information for the scientists and managers of an organization. The requirements of external accountability often motivate managers to improve internal management processes that will facilitate the fulfillment of these requirements. When an agricultural research organization strengthens its accountability with regard to farmers and industry this has a direct impact on management decision making and overall performance.

Public accountability

Evaluation of agricultural research activities is in many cases motivated from outside the institution —by a donor organization, the national treasury, or groups of producers. Researchers rarely initiate their own evaluations to improve their work . Often, when external pressure is removed (for example, when a project funded by a donor ends), monitoring and evaluation also are stopped.

Information for planning and management

One of the most positive achievements in recent years has been the growing acceptance that evaluation can play a useful role in agricultural research management. A similar evolution has been observed in other fields, such as in education, where evaluation was first considered as a way of complying with external requirements for impact information, then as a way of measuring whether a program's achievements met the proposed goals, and finally as a means of helping management improve program planning and management.

Experience of the donors

Donor agencies evaluate their projects mainly for reasons of accountability. The most common criteria of a donor's evaluation are the timely and appropriate use of funds, and the fulfillment of planned objectives. Many agencies apply a management tool known as the logical framework (presented in Manual 3) for planning, monitoring, and evaluating projects. Development banks also use estimates of the rate of return to their

investments. Some agencies evaluate the impact of specific projects (generally the most successful). Other donors have promoted periodic evaluations for improving project management and developing the local evaluation capacity (Figure 9).

Meta-evaluation

If the evaluators are going to practice what they preach, evaluations should also be submitted for evaluation, along with the other activities of an agricultural research agency. Evaluators call this meta-evaluation. In general, the emphasis should be on having achieved the purposes of the evaluation, and on the usefulness of results. This means that meta-evaluation should emphasize the practical aspects and the usefulness of results rather than the rigor and precision of the methods used to collect and analyze information.

As in other types of evaluation, meta-evaluation can be done by people within or outside the organization, by evaluators, professionals, or by specialists in various topics or disciplines; it can be done either by individuals or evaluation teams. At a minimum, evaluators should do a "self-assessment" of their work, in order to learn from the experience and improve future evaluations.

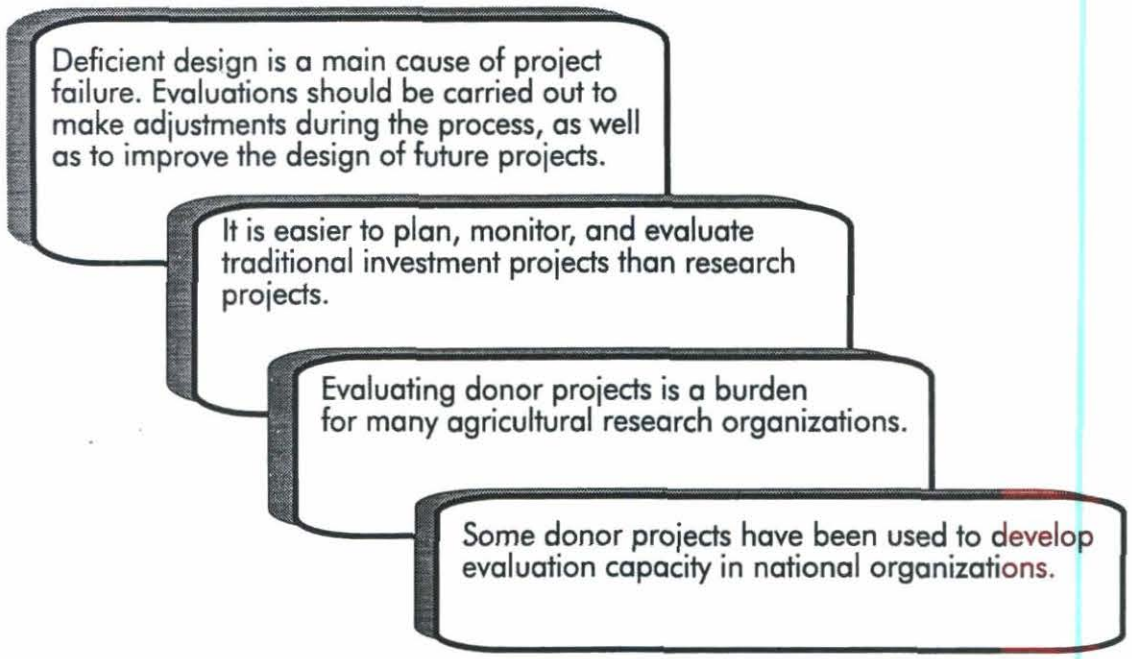


Figure 9. Lessons learned from donor evaluations.

Institutionalizing Evaluation

Periodicity

Management of an evaluation includes the articulation and supervision of a series of activities, such as setting goals, collecting and analyzing information, presenting the report, and supervising the implementation of recommendations.

In research, it is difficult to clearly define phases and predict results. Activities tend to be continuous, over long periods of time. In fact, instead of ending, they tend to evolve into other activities. Progress can vary according to resource allocation, availability of staff, and competitive activities.

Management should organize annual internal reviews of projects, programs, and research station operations. External reviews and planning exercises are useful in cycles of three to five years. External evaluations can also be useful when there is a major change in direction, funding, or the mandate. Impact studies can be appropriate when performance and benefits of the research organization are questioned by those who fund research (INTA, 1991).

Many options are available for organizing and managing evaluations in an agricultural research organization. It can be done from the office of the director of the institution or of research. It can depend directly on the governing board. It can be managed by a planning and evaluation unit or by a specialized evaluation unit (independent of planning). The procedures can be very centralized (with all the guidelines from headquarters) or they can be decentralized. by centers and research stations. An external review can be commissioned by the same organization or by its sponsors. Emphasis can be given to the uniformity of evaluation procedures and quality control, or to carrying out flexible evaluations that satisfy the specific needs of management and staff in special situations.

Institutional structure

Institutionalization experiences and alternatives

No universal rule is valid for structuring and managing an evaluation. Evaluation systems and procedures need to adjust to the structure and culture of the specific organizations. They also need to evolve as the institution itself changes. For example, in a highly centralized organization, evaluation processes are also bound be centralized. But, if decentralization begins, the evaluation system should also be decentralized to bring decision making and accountability closer to where research activities take place.

Institutionalizing evaluation requires at least three conditions:

- Interest in and support for evaluation on the part of management and staff,
- Clearly defined units of research (e.g. projects or programs)
- Capacity to handle the information generated by the evaluation (Figure 10).

To be efficient, an organization's evaluation system must be linked to a broader, functioning information system that allows information to arrive selectively and at the proper time to the different decision-making levels, from researchers to directors.

A favorable attitude toward evaluation, training in project preparation and management, and the support of a good computer system all prepare the way for institutionalizing evaluation in an agricultural research organization. This process is expensive and requires time, but if implemented correctly, makes the organization more efficient and effective.

The Latin American and Caribbean Seminar on Mechanisms of Evaluation in Agricultural Research Institutions, which met in Paipa, Colombia in 1988, recommended a model for institutionalizing evaluation (Novoa, 1989). The model considers the operational and structural means, as shown below.

Operational means. *The proposed model emphasized participation at every level, including as far as possible technical staff, producers, unions, directors, extension agents, and consumers.*

The most common means of evaluation emphasized were meetings and periodic reports at the level of operational units. Their frequency should be kept low so researchers are not distracted from their fundamental tasks; but they should be carried out at least once a year.

Review and external evaluation missions, commonly used when programs or projects are financed with external funds or credit, are useful evaluation mechanisms when they combine personal and internal criteria, and when the terms of reference are agreed upon by both parties.

National entities should have a technically-based evaluation system to satisfy all evaluation interests and needs, with a minimal combination of external elements.

Research institutions need to initiate periodic external evaluations (every 3 or 5 years) to support management.

Institutions also need to evaluate their budget exercise and human resources as a complement to the evaluation of objectives, results, and effects of research. The use of financial and human resources, and incentives should be taken into account.

Structural means. *In order to put the described evaluation mechanisms into practice, a central co-ordinating unit should be established, responsible for consolidating institutional evaluation efforts setting guidelines, and implementing their execution (developing formats and manuals if necessary). This unit could also be in charge of planning, monitoring, and evaluation in the institution.*

This unit would receive, compiles, and synthesize evaluation reports prepared by the operational units, thus consolidating a global institutional evaluation process. Through the respective analysis the unit would carry out the evaluation: At the global level, it would measure achievement of institutional objectives. Given its role, it serves as a management support unit, with a small staff requirement.

Management of the evaluation process

Whatever the type of evaluation, it should be managed to minimize the disturbance of scientific work: It should also be organized to provide information to the planning, budgeting, and staff management cycles. Whoever acts as the manager of any evaluation process should become the "facilitator" so that the different activities will be carried out appropriately.

Evaluation capacity

The attitude toward evaluation in an organization determines, to a large extent, whether an evaluation should be internal or external. A combination of internal and external evaluators is recommended if the organization has a tradition of evaluation and recognizes its benefits. The aspects to be evaluated must be analyzed beforehand, in order to choose the appropriate type of evaluation. Experience shows that institutional evaluation structures are generally created in response to strong pressures from top management or external sources (donors, the public treasury, etc.). Researchers rarely promote the evaluation of their activities; they are more likely to reject the idea. Recently agricultural research institutions have suffered from their countries' structural programs, and have been forced to give more precise and timely information to ensure their continuation either with regard to the funding they receive, or simply the credibility of their activities. With these new trends, institutions have, in many cases, formed teams responsible for evaluating their technical activities.

These evaluation teams have prepared evaluation proposals, designed instruments for collecting data, and organized different types of evaluations. They have also enlisted the participation of researchers in the evaluation process. The degree of staff participation, the scope and quality of evaluation work and the development of information systems varies widely from institution to institution.

Desirable characteristics of an evaluation manager

Experience shows that the evaluation manager should have certain personality traits and skills, such as leadership, sense of opportunity, and flexibility. These characteristics are essential for managing an evaluation process suitable to the characteristics, conditions, and resources of an institution. His common sense should make him aware of the training and adjustment needed to implement the system. He should also guide the different phases of organizing, implementing, and consolidating the institutional evaluation system.

The case studies and the authors' experience indicate that the process of institutionalizing research evaluation has begun only recently, although many managers are attempting improve evaluations, based on their own practical experience, plus their knowledge of experiences of donor agencies and the industrial countries.

Characteristics required of the evaluator

Choosing the evaluators is a crucial operation, because an evaluation's success depends on their credibility and their skill in carrying out the evaluation.

Experts hired from outside the organization are generally more objective. They may also be chosen for special skills or external perspectives. As pointed out earlier, using external evaluators has three main disadvantages: 1) they may fail to understand key aspects of the organization's culture; 2) the knowledge and experience they acquire in the course of the evaluation are lost to the institute when they finish their job; and 3) their recommendations may not be adequately implemented if formal monitoring mechanisms are not established for the use of the results. These disadvantages can be partly resolved by giving evaluators adequate information about the institution well in advance, before they begin their task.

Training

It is desirable to provide training for those responsible of managing an evaluation system, as well as for research managers in the institution.

Training in evaluation should preferably be programmed periodically through workshops, to favor the interchange of experience and to

Figure 11. The evaluation co-ordinator

overcome difficulties caused by a lack of understanding of the evaluation process, its management, applicability, and institutional benefit.

Resources for the evaluation

Evaluation design should be realistic about the following aspects: availability of resources (such as for staff, consultants, and travel); preparation, production, and distribution of reports; workshops and meetings; data processing; supplies and materials; and communications. The resources available should be compared with those needed for different evaluation approaches. Once this is done, management can decide which resources to invest, what methods to use, and what products to expect.

Direct and the indirect costs should be considered in establishing an evaluation. **Direct costs** are influenced by the type and the quantity of information that will be collected, where it will be collected (in the research station or on remote sites), and the cost of the staff. Impact evaluations, which gather a large amount of information from households and farms and use sophisticated analyses, tend to be more expensive.

Indirect costs—which occur when personnel are separated from their main research function—can be significant, and may be even higher than direct costs. For example, preparing for an external review can be a useful exercise that helps identify weaknesses and adopt corrective measures. But the indirect cost of staff preparations may distract them seriously from their research work and exceed the direct cost of contracting the evaluation team.

Programming the evaluation

The general rule is that the responsibility for programming the evaluation should be associated with top-level decision makers, for example the institution's directors and the board, and not the administrative departments. This gives an evaluation more credibility, and stimulates the actions needed in response to evaluation findings (INTA, 1991).

A checklist is useful for designing an evaluation. It is also useful to plan a list of specific activities, indicating when each should be carried out, who will be responsible, and how to carry them out. Table 10 presents a checklist for designing an evaluation.

In programming an evaluation, it is important to establish deadlines for collecting and analyzing information, and for writing reports. It is a common error to spend too much time collecting information, and too little time designing the evaluation, doing the analysis, and preparing the report.

Table 10. Checklist for designing an evaluation

- Define the subject:
  What needs to be evaluated?
- Identify the client:
  Who is the evaluation being done for?
- Clarify the purpose:
  Why is the evaluation being done?
- Identify the issues:
  What questions need to be answer?
- Assess the resources:
  What resources are needed or available?
- Organize the evaluation:
  Who, within the organization, should be responsible for the evaluation?
- Select the evaluators:
  Should the work be done by the organization's staff or by outsiders (professional evaluators, subject matter specialists, or clients)?
- Select the methods:
  How should the information be collected and analyzed?
- Decide on reporting:
  When should reports be made and to what audiences?
  What should be their content and style?
- Determine follow-up:
  What follow-up should be made after the evaluation?
  Who should monitor the follow-up?

Source: Horton *et al*, 1993.

Conclusion

The best way to integrate so many interacting actors, factors, and actions in an evaluation process is through the CIPP approach. Figure 12 presents a model whose **objective** is to organize an evaluation system and whose **end** is to change the institution's course and guarantee its continuity.

Summary

One important aspect of institutionalizing evaluation is periodicity: management cycles should be established for setting (and reviewing) goals, for collecting and analyzing information on the implementation of research and on progress toward goals, and for taking stock and evaluating research

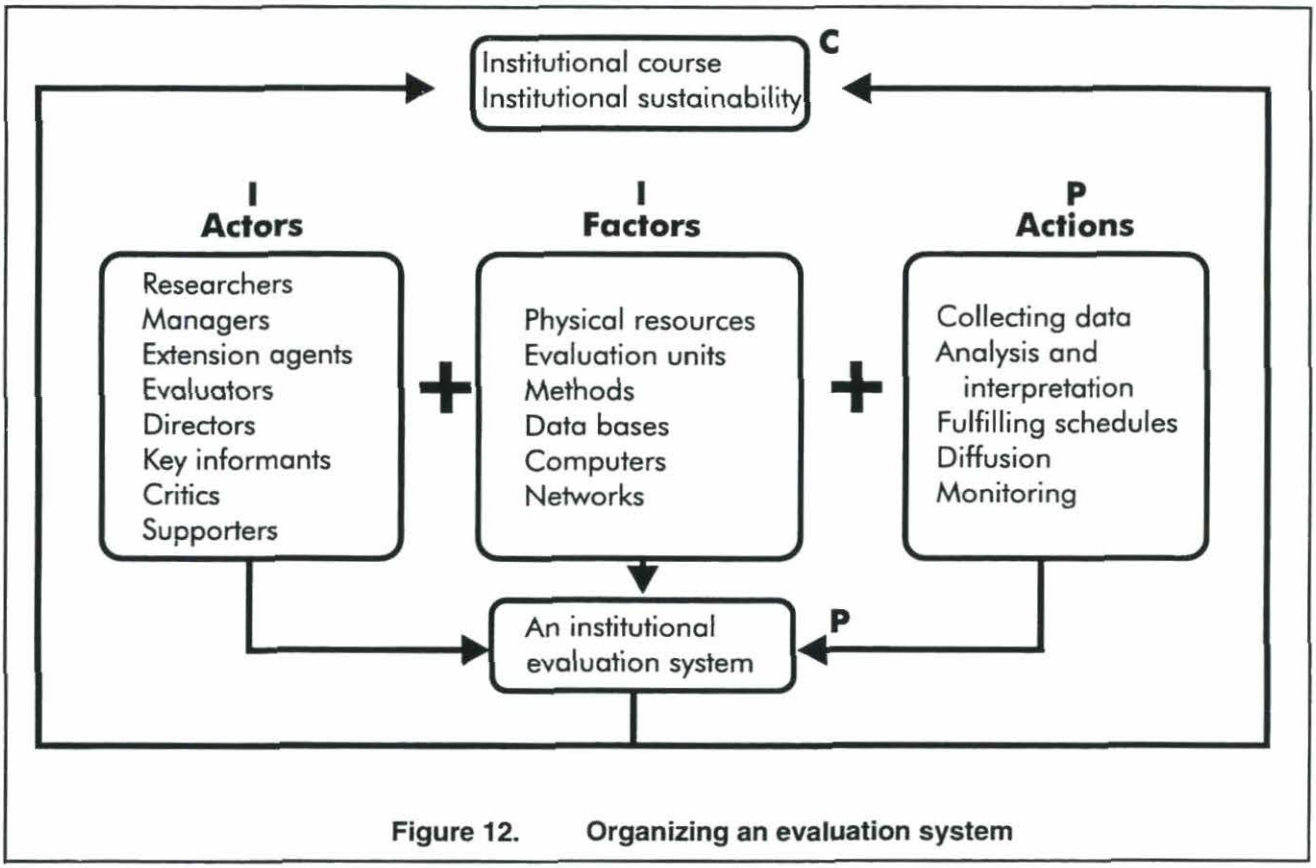


Figure 12. Organizing an evaluation system

work and results. In many instances, two interrelated cycles are useful: an annual cycle and a three-to-five year cycle. In each cycle, the role of evaluation, its procedures, and the use of its results need to be clearly defined. Within such a cyclical dynamic, the evaluation system should be flexible enough to respond to unexpected events such as a change in top management, which might request a special review or evaluation exercise.

Institutionalization of evaluation cannot be done over night; it involves a process. Such a process requires leadership and technical competence. Experience indicates that the process can be effectively led, organized and managed by a small technical team.

In many cases it is advisable to create an evaluation unit, or PM&E Unit. During the process of institutionalization it is important to define the role of the evaluation unit or group, and to make its role known to all staff members. As a rule, the evaluation unit should not carry out all of the organization's evaluations, but to support and coordinate the evaluation process, provide methodological guidelines, and prepare synthesis evaluation reports based on the evaluations that are carried out.

Both human and financial resources are needed for evaluation, and these should be anticipated. In most cases, specialized training will be needed in evaluation —both for the technical team and for

managers and staff more broadly. Direct (operational) and indirect (staff-related) costs of evaluations should also be anticipated, to ensure that the resources needed will be available and that costs are reasonable and feasible.

An institutional evaluation process needs to be designed by, or with, top and middle management, or else it runs the risk of not being accepted or implemented fully.

Before implementing an evaluation process, it is advisable to plan it carefully as an institutional project with a clear goal, purpose, expected outputs, activities and resources required. Such a project statement, should spell out those responsible for different evaluation activities, the reports needed and their due dates, the types of evaluations to be carried out, and the resources and training needed.

Appendices

Page

Appendix 1. Terms Used in the PM&E Manuals 52

Appendix 2. Summary of the Types of Evaluations Used in Research Organizations 56

Appendix 3. Individual and Collective Variables 57

Appendix 4. Evaluation Scale for Interviewers 58

Appendix 5. Planning for Computerized Analysis and Data Processing 58

Appendix 6. Planning, Monitoring, and Evaluation in the National Institute for Agricultural Technology (INTA), Argentina 59

Appendix 7. Bibliography 66

Appendix 1. Terms Used in the PM&E Manuals

The training materials on PM&E use a number of general concepts related to agricultural research management. Not strictly limited to definitions of terms, they propose concepts that reflect the thinking of the authors in relation to the general theme.

Accountability
The obligation to report, explain, or justify something. The responsibility of an organization or its staff to provide evidence of research expenditures and performance to donors or higher levels of management.

Assumption
A fact or statement that is accepted as true. In relation to the logical framework, it is a statement about factors that can influence the achievement of objectives but which are beyond the control of researchers, such as political or economic policies or the availability of farming inputs.

Beneficiaries
People, households, organizations, communities, or other units that are affected positively by (or *benefit* from) a research program or activity.

CIPP evaluation model
A conceptual framework for improvement-oriented evaluation. CIPP stands for four kinds of evaluation:

- *Context evaluation.* Assessing the context of a program, identifying target populations and their needs, identifying opportunities and problems in addressing needs, and judging the responsiveness of goals and objectives to assessed needs.
- *Input evaluation.* Identifying and assessing alternative strategies, schedules, budgets, resource requirements, and procedural designs needed to accomplish the goals and objectives of a research activity.
- *Process evaluation.* Assessing the implementation of a plan by recording and judging ongoing activities and accomplishments in relation to the procedural design. It provides information helpful for changing operational plans during implementation.
- *Product evaluation.* Measuring, interpreting, and judging the attainments of a research activity. Intended to interpret the work and merit of an activity's final outcomes in relation to the needs of the group it is intended to serve.

Clients
The intended users of agricultural research products, generally including farmers, agribusiness entrepreneurs, policymakers, extensionists, and consumers.

Criteria
A standard of judgement. The basis for a comparison, a test or an evaluation.

Decision-making level
The level within a research organization or system (for example, the level of the researcher, project manager, experiment station or institute manager, or policymaker) at which a particular decision is made, or to which an evaluator reports.

Effectiveness
The degree to which an activity, project, or program attains its objectives. The extent to which outputs are obtained and effects achieved in relation to objectives.

Efficiency
The degree to which an activity produces outputs at the least cost.

Evaluation
Judging, appraising, or determining the worth, value, or quality of research — whether it is proposed, ongoing, or completed — in terms of its relevance, effectiveness, efficiency, and impact.

Ex ante evaluation
An assessment done before research begins, usually in terms of its relevance, feasibility, potential impact, or expected benefits. Can be used to define a baseline against which progress towards objectives can be measured or to set priorities among several research areas.

Expert review
(See *peer review.*)

Ex post evaluation
An assessment of an activity or its outputs after the activity has been completed. The purpose is usually to estimate benefits in relation to costs.

External analysis
Sometimes called prospective analysis of the external environment (or context analysis). The process of assessing and evaluating the external environment, to identify present and potential opportunities and threats, which can influence the institution's ability to achieve its objectives. (See also *organizational analysis.*)

External environment
In the case of agricultural research the macro-environment that affects an institution, program, or project. At this level, events are practically beyond the organization's control. Examples are governmental policies, consumption trends and development of new scientific knowledge.

External review
Evaluation of a research system, organization, program, or project carried out by persons from outside the unit being evaluated. Usually conducted by experts or peers, but research clients, supporters, or stakeholders may also participate in the evaluation.

External validation
The process by which internal decisions are discussed within external stakeholders, in order to confirm or revise them. In strategic planning, conclusions about threats and opportunities, and the mission, objectives, and policies are generally validated externally.

Formative evaluation
An evaluation aimed at providing information to planners and implementors on how to improve an ongoing program or project.

Gap analysis
An assessment of the requirements of a research plan in terms of the resources needed (financial, human, and physical) to achieve the desired goals.

Goal
Used in the logical framework, a goal is the ultimate end or objective towards which a research activity, project, or program is directed. It is usually something like improving incomes for farmers. (See also *objective*, *purpose* and *output.*)

Impact
The broad, long-term effects resulting from research, usually economic, social, and environmental.

Input
In terms of the logical framework, inputs refer to the resources needed to implement a project, including personnel, operating funds, facilities, and management.

Institutional sustainability
An organization's condition of being accepted and considered legitimate by society. Institutional sustainability has several requirements including (a) an institutional project (clearly defined mission, objectives, policies, and strategies); (b) institutional competence; (c) institutional credibility.

Institutionalization
A process that impersonally establishes a structure, plan, program, project, or activity in the day-to-day operation of an organization.

Internal review
Evaluation of a research project, program, or organization that is organized and carried out by the management and staff of the unit. (See also internal program review).

Logical framework
Often called the *logframe,* it is a tool for planning, monitoring, and evaluating projects in the broader context of programs and national goals. It clarifies the logical links between project inputs and a hierarchy of objectives: direct outputs, broader purposes, and the ultimate goal.

Means of verification
The sources and methods used to obtain and assess information about the achievement of research objectives.

Metaevaluation
Critical assessment and overview of evaluation procedures and experiences. Metaevaluation is done to learn from past evaluations and improve future ones.

Mission
The official statement of the reason for an organization's existence — its basic goals and purpose. (See also *strategic planning.*)

Objective
The expected output, purpose, or goal of a research effort; something towards which efforts are directed. Objectives may also be specific operational statements regarding the desired accomplishments of an activity. (See also *goal*, *output* and *purpose.*)

Objectively verifiable indicator
Specific measures of progress or results at a specific level of a project's hierarchy of objectives.

Ongoing evaluation
Evaluation carried out during implementation of an activity. It involves observing or checking on research activities and their context, results, and impact. Ensures that inputs, work schedules, and outputs are proceeding according to plan (in other words, that implementation is on course). It also provides a record of input use, activities, and results and warns of deviations from initial goals and expected outcomes. (See also *monitoring.*)

Operational planning
A process for defining what an organization intends to accomplish, how and when this will take place, and who will be held accountable.

Organizational analysis
Internal analysis carried out by gathering and assessing information on the inputs, processes, and products of an organization. The purpose is to identify strengths and weaknesses in relation to opportunities and threats posed by the external environment, and in relation to the organization's objectives.

Output
The specific product or service that an activity produces or is expected to produce. Used in the logical framework to refer to specific results for which the project manager may be held accountable, such as the release of a new maize variety. See also *goal*, *purpose* and *objective.*

Participatory management
Creating a culture of effective participation of an organization's members at all levels. It involves sharing ideas and responsibilities, and getting members' commitment to design and carry out activities that will contribute to institutional objectives and bring about desired institutional changes.

Peer review
Process by which the scientific merit (conceptual and technical soundness) of a research proposal, publication, or activity is evaluated by other scientists working in the same or a closely related field.

Planning
A process for setting organizational goals and establishing the resources needed to achieve them. It is also a way of building a consensus around the mandate, direction, and priorities of a research program or organization.

Policies
Major guidelines for reaching ends in accordance with priorities. Policies should be formulated after, or as a consequence of, the formulation of the organization's mission and objectives. Policies give direction to decisions on inputs and processes.

Products
Specific goods or services produced by an organization program, project or activity. (See also *outputs.*

Program
An organized set of research projects or activities that are oriented towards the attainment of common set of objectives. A program is not time-bound, as projects are, and programs are higher in the research hierarchy than projects.

Programming levels
The areas that encompass activities of an agricultural research institution, according to the *specificity* of the objectives. The two most common levels are projects and programs.

Project
A set of research activities designed to achieve specific objectives within a specified period of time. A research project is composed of a group of interrelated research activities or experiments that share a rationale, objectives, plan of action, schedule for completion, budget, inputs, outputs, and intended beneficiaries.

Project cycle
A framework for planning and managing projects. It is composed of distinct phases through which a project moves during its lifetime. Variations of the project cycle are used to manage large-scale investments, development-agency activities, and various kinds of research.

Project management
A framework for the systematic planning, implementation, monitoring, and evaluation of research projects and activities.

Purpose
The desired effect or impact of a project. (See also *goal, output,* and *objective.*)

Quality control
A set of planned and systematized activities to guarantee that the products and services of an institution will fulfill the expectations of the public, beneficiaries, and stakeholders.

Relevance
The appropriateness and importance of research activity's objectives in relation to broader (e.g. regional or national) goals or clients' needs.

Scenario
The simulation of a probable future situation, in the context of the institution's location, taking into consideration the interaction among economic, political, social, and cultural factors, and how these may affect the institution's ability to act.

Stakeholders
Groups whose interests are affected by research activities. The stakeholders of a research organization include staff members, farmers, and extension agents, among others.

Strategic planning
A process by which an organization builds a vision of its future and develops the necessary structure, resources, procedures, and operations to achieve it. The process is generally participatory, and based on analyses of the external environment, the organization, and "gaps". External opportunities and threats and internal strengths and weaknesses are assessed. This is followed by formulation of the organization's mission, objectives, policies, and strategies. Strategic planning is long-term in nature (e.g. for 10 or more years.) It serves as a base for tactical and operation planning. (See also *tactical planning* and *operational planning.*)

Strategy
A course of action involving a logical combination of actors, factors and actions chosen to reach a long-term goal or vision. It is important to distinguish policy from strategy. Policies are general guidelines to achieve given objectives. In addition, Strategies incorporate a logical sequence of steps. (See also *strategic planning.*)

Summative evaluation
A summary statement about the accomplishments, effectiveness, value, and impact of programs. Summative evaluations are made for accountability purposes and for policy-making.

Survey
A technique for gathering information from individuals or groups. It can be done by observing, administering questionnaires to, or having discussions with members of the group being surveyed.

Tactical planning
A process of organizational planning at the intermediate management level. The objectives, goals, policies, priorities, and strategies defined through tactical planning are for the medium term (generally 3-5 years); they are based on the strategic planning, and are the guidelines for the operational planning.

Appendix 2. Summary of the Types of Evaluations Used in Research Organizations

Types of Evaluation	CARDI	SIRI Jam	CONITTA CR	MAG CR	ICTA Guat.	INIFAP Méx.	CIAT Bol.	ICA Col.	CENICAFE Col.	INTA Arg.	EMBRAPA Bra.	INIA Chile	Ag-Can Can	ARC USA	TOTAL
External review															
Organization	1	0	1/4	0	1/2	0	1/4	1/2	1	1/2	1	0	0	0	5
Program	0	0	0	0	0	1/2	1/2	1/2	1	0	0	0	1	1	4 1/2
Unit	0	0	0	0	0	0	0	0	1	1	1	0	1	1	5
Project	1/2	0	0	0	1/2	1/2	1/2	1/2	1	1	1	1	1	1	8 1/2
Internal review															
Organization	1/2	0	1/2	0	1	1/4	1/2	0	0	0	0	0	0	0	2 3/4
Program	1/2	0	1/2	0	1	1/2	0	1/2	1	1/2	1	1	1	1	8 1/2
Unit	0	0	0	1/2	1	1/4	0	1	1	1	0	1/2	1	1	7 1/4
Economic evaluation															
Program	0	0	0	0	0	0	1/2	0	0	1/2	1/2	1/2	1/2	1/2	3
Project	1/2	0	0	0	1/2	1/4	0	1/2	0	1/2	1/2	1/2	1/2	1/2	4 1/4

1 = Has systematically developed reviews/evaluations and uses them regularly
½ = Has developed procedures and uses them on an ad hoc basis
¼ = Has limited experience in review/evaluation procedures
0 = Has not developed evaluation procedures
Source: Uribe and Horton, 1993.

Appendix 3. Individual and Collective Variables

Variables may be classified as individual and collective variables (Lazarsfeld and Menzel, 1969).

Individual variables

Individual variables, properties that characterize individuals, include the following sub-types:

Absolute variables characterize individuals without needing to refer to a property or characteristic of a group. Examples are age, occupation, income, and education level.

Relational variables are obtained from information about relationships among members of a group. The "popularity" of a person, for example, can be defined operationally in terms of the number of positive references given by his or her colleagues.

Comparative variables are properties that characterize people with reference to a certain value given in the group. Thus, each member of a group can be identified as being older, or younger, or having the same age as the average age for the group. The property or comparative variable in this case is age.

Contextual variables are collective properties used to characterize people. If a region is known, for example, for having a high degree of illiteracy, this situation or property may be used to characterize the people of this region, (taking care not to imply that all are illiterate!)

Contextual variables were used by the French sociologist Durkheim; in a study on suicide, he showed that the rates varied according to the social-cultural context.

Collective variables

Collective variables refer to properties of groups, and are divided into the following sub-types:

Analytic variables are properties obtained by carrying out a mathematical or statistic operation on a feature found in each and every one of the units making up a group. For example, the average age of a group of people, the percentage of illiteracy, belong to this type of variables.

Structural variables are obtained by carrying out operations with the data obtained from members of a group that have interactions or social relationships among them. For example, the cohesion of a group can be defined as the proportion of "sociometric choices" made within the group.

Global variables are properties that characterize the group without referring to properties of the individual members. For example, whether a neighborhood council exists, or a hospital or a school, etc., in a region, constitutes the global properties of that region (Briones, 1982).

Appendix 4. Evaluation Scale for Interviewers

Interview			
	Scale		
The interviewer	G	A	D
• Had the materials and guide ready for the interview			
• Began the interview at the time agreed			
• In the course of the interview...			
• Greeted the person to be interviewed cordially			
• Made sure the person interviewed was comfortable			
• Broke the ice			
• Explained the objectives and components of the interview			
• Obtained the information required			
• Avoided getting off the main subject			
• Allowed the person interviewed to express him/herself freely			
• Completed the interview in the time assigned			
Attitudes			
• Showed diplomacy and courtesy			
• Was always alert			
• Developed empathy with the person interviewed			
Observations _____			
_____			
_____			
Signature of evaluator:_____Signature of person evaluated: _____			

G = good; A = acceptable; D = deficient

Source: Zapata, S. V. 1995.

Appendix 5. Planning for Computerized Analysis and Data Processing

Nowdays, computers are commonly used for processing and analyzing research and evaluation data. The computer's advantages include its speed and the volume of data it can handle. Only a few minutes are needed to process huge amounts of data, the analysis of which would take weeks or months if done with a desk calculator.

The economy and efficiency of a computer depend, however, on certain practical aspects:

1) The **information must be coded** so that it can be entered into the computer. Generally, this presents format restrictions in that all entries receive alphanumerical values (that is, letters or numbers). Someone must enter the data correctly following the logic of a specific set of codified instructions.
2) A **computer *program*** processes and analyzes data. If you can adjust your design and analysis needs to *existing* programs, you will not incur additional costs for writing up a new program or

modifying old ones. Therefore, it is important to find out if existing computer services and programs are adapted to your specific needs before planning the computerized analysis of your data.

When planning to process data in a computer, if you are not familiar with the computer, you should ask a computer technician to advise you on how to code and process the data. In general, each unit of information should be coded to represent one or more "columns" that correspond to that entry's "row" in a matrix. Run a control test of recording and processing procedures to be sure you can generate results you expected, and to solve any problems that might come up.

Some precautions on computerized analysis of research data

The power and the prestige of modern computers can fool the unsuspecting researcher. Take heed of the following:

1. **Errors.** A computer's mechanisms and electronic circuits tend to be highly trustworthy but humans make many mistakes! For example, data can be coded or entered incorrectly; there an error in the computer program; the specific instructions for a particular program can have errors; what is more, the magnetic tapes or diskettes used to store programs or data may be damaged. For these reasons, entries need to be checked, and computer programs and procedures need to be carefully tested.
2. **The problem of the "black box."** For a many researchers, technical details having to do with the computer, its programs, and its statistical manipulation are a sort of "black box." You cannot see what goes on inside and accept blindly what the computer prints out, as well as computer experts' opinions on the whole mysterious process. For some research purposes, this does not create a problem. However, in the evaluation process, the scientist or technician may want to stay in contact with the data, "feel close to them". This sometimes means processing data by hand. One of the authors of this module generally compares the computerized analysis of his data with a manual analysis of a sub-sample. In this way he keeps in contact with his data and sometimes finds errors in the computer analysis (Isaac and Michael, 1974).

Appendix6. Planning, Monitoring, and Evaluation in the National Institute for Agricultural Technology (INTA), Argentina[1]

From its establishment in 1956 until 1986, PM&E activities at INTA were handled by what was then called the National Service for Technical Programming and Evaluation. Later they were assigned to the National Direction of Planning and Evaluation This centralized system was inadequate for providing adequate and timely information and decisions. This may have been due to the fact that a centralized PM&E system cannot be very effective in a research system as vast as that of INTA and with such a wide geographic coverage.

Following a new institutional plan which included as one of its objectives political and operational decentralization, the National Deputy Directorate of Planning and the National Deputy Directorate for Control and Evaluation (NDDC&E) were established in 1987.

NDD C&E has organized a system for the monitoring and evaluation of technical activities, based on the active participation of INTA's 18 research centers. The Deputy Directorate has acted as organizer, coordinator, and promoter of the M&E process.

From 1987 to 1989, the information, monitoring, and evaluation systems were organized under INTA's new decentralized scheme. An "Institutional Evaluation Program" was initiated in mid-1990.

Monitoring

Information system

Before analyzing the types of monitoring carried out at INTA, it should be emphasized that INTA's new PM&E system is based on the premise that management of technical activities requires an efficient and transparent information system with precise and timely information. Thus, the information system is the basis of the whole PM&E process.

Source: Hogg, D. K. 1994
Note: This document was written in August, 1993 and the information refers to this date.

All of INTA's technical activities should be reported, so that all levels of the institution are aware of them. Information and analysis are used within a permanent process of monitoring, critical analysis, and subsequent evaluation.

The following proposals and reporting documents contribute to monitoring and evaluation at INTA:

Proposals plans	Reports
National Agricultural Technology Plan ("PLANTA")	INTA's Annual Proceedings
Regional Technology Plan ("PLANTAR")	Annual reports of research and experiment stations
National programs and subprograms	Annual program and subprogram reports
Projects* outlines	Annual and final project reports
Research workplans*	Annual and final reports on research experiments
List of summary information research experiments workplans	List of summary information information on technical activities

* The 2 basic unit of work at INTA are "projects" and "workplans". A project is generally team efforts designed to solve a specific problem in the field; as such , it may involve both research and extension activities. A workplan is generally for a more narrow research activity.

Components of the control and evaluation system

Because workplans are the simplest instruments for planning and monitoring, the following are mechanisms used to present, monitor, and communicate annual or final reports on those plans.

Work plans are proposed by researchers, or extension technicians, based on a standard model that analyzes the completition of proposed activities, the state of progress, achievements, and results.

INTA's methodology for presenting technical proposals allows these to be analyzed quickly. It includes elements of the logical framework to facilitate monitoring and evaluation. Just as specific instruments are used for presenting and monitoring workplans, other instruments are available for projects.

Data bases. The data base of the control and evaluation system is an essential information instrument for monitoring technical activities at INTA. About 1300 workplans make up the database.

Information keyed into the data base can be used to generate a summary information list of technical activities. This information, completed for each approved project and workplan, is sent to the National Deputy Directorate for Control and Evaluation at headquarters. Data is also incorporated into local databases in the center or experiment station where the workplan or project originated

The information includes a code number for each workplan and project, assigned by the Center; the title; location by center and unit or rural extension agency, program, subprogram; mandate (national or regional programs); type of research (basic, applied or adaptative); and name of the person responsible for the work. Information on the annual budget, other participating institutions, and types of support is also included.

Finally, the information includes a summary describing the initial problem status, the objectives and methodology to be used, and the final situation to be achieved. "Agrovoc" descriptors and key words are provided to widen the possibilities for searching for and classifying the information.

All information on INTA's technical activities existing in this database can be accessed from any personal computer at any of the experiment stations in the country.

This data base can be linked through the code number of each project and with the database of the budget system and resource use, which gives access to multiple information on resources used by program, subprogram, region, product, or staff member.

To sum up, the databases give access to project and work plans, and can generate much useful information for monitoring tasks. They are useful for the institution's professionals who have access to them through their personal computers, so that they can follow the progress of the work that interests them. Anyone responsible for monitoring also has access to the databases.

The data base for monitoring was developed by the NDDC&E, and was completed with the participation of all Centers. It allows multiple outputs. For progress indicators it uses data from the workplan reports; these enable comparisions to be made between the initial situation and the proposed final objective, which are in the same base.

Information in the data base and its multiple outputs have been useful as a source of information for the whole process of monitoring, for internal and external evaluations, and for general information related to resource use in technical activities at INTA.

Kinds of monitoring

The database was organized in a decentralized way, following simple guidelines. The first obstacle was training the staff to operate the computer equipment, but this was overcome rapidly at the Centers. The limited data processing capacity of most of INTA's computers slows down information retrieval, but these problems could easily be technically resolved.

Monitoring is carried out formally during the Annual Internal Evaluation of each Center. Work, programs, and regional projects plans are supervised during these evaluations. This fulfils two aims. Firstly, to inform those responsible for the units (unit directors, program coordinators) of the state of progress of technical activities. Secondly, to inform those responsible for INTA policies (Local Councils and Center Councils, experiment station advisors) so that they can compare the activities carried out in pursuance of the guidelines laid down by the regional plans.

Internal evaluations at INTA are organized around the general guidelines set down by NDDC&E, and by any proposals, resolutions, or specific dispositions generated by the different Centers in reference to the activity. Internal evaluations are generally annual; participants include members of the Center Council and the directors of Centers and Units, Area co-ordinators, and the technicians responsible for the specific tasks.

National Program and Subprogram Co-ordinators serve as consultants during internal evaluation. Their participation is essential when considering specific aspects of the different subprograms. NDD C&E provides conceptual and methodological support.

Methods and techniques

Papers are presented at meetings with a wide participation and exchange of opinions, and in groups for analysis, conclusions, and elaboration of proposals. Defined formats are used for presenting work proposals and annual reports to provide information on activities and products for a given period. A summary of what was done during the period, in terms of objectives, methodology, and results is recorded in one of these formats. The plan and project work proposals are also contrasted with the Center's technology plan.

The monitoring and evaluation of regional projects is done according to guidelines defined by each Center, within the framework of general guidelines given by NDDC&E.

A formal report form the Center is prepared every time that partial or complete evaluations are carried out. This documents the tasks of monitoring and evaluating plans and projects and serves as a reference for future policy or operational decisions.

Regional and Research Centers prepare an annual chronogram of monitoring and internal evaluation activities which they submit to National Program co-ordinators and NDDC&E to prepare them for annual monitoring and internal evaluation reviews.

Since the Institutional Evaluation Program was put into operation, numerous internal evaluations have been carried out at INTA's Centers. In 1991, 16 of the 18 Centers completed formal monitoring activities. In 1992 all 18 Centers submitted their monitoring chronograms and the programming of those activities.

The progress indicators used for monitoring are set previously by plans or projects both at product and at activity level. These indicators must be defined in the presentation documents or in the annual reports.

INTA has a set of standard procedures for elaborating budgets for work plans and projects. Work plans should be accompanied by a budget when they are presented to the Center Boards for consideration and approval. If they are approved, the resources are allocated, as part of the Center's budget. Once a year, the plans compete for resources within in the Center. The allocations are communicated to the Central Management unit; every month this unit sends funds for the approved plans to the corresponding Center, plus correction for inflation.

The Central Administrative Directorate has a data base with information on the budget and the total amount spent by each plan; the code number in this data base corresponds to the code in the Control and Evaluation data base. Thus information on economic resources is easily available on, for example, resources used by a plan on, by the plans that integrate a program, or by a Center's programs, etc. With this information, the use of resources can be compared among regions, programs, items, disciplines, researchers or extension agents, and can even be compared by type of expenses, such as travel expenses, transportation, agrochemicals, and others.

The project budgets are prepared following the norms of the "Guide for presenting and monitoring projects," which includes an extremely detailed list of yearly activities to be budgeted.

Once a project is approved, the executive unit receives funds every month for that month's activities. Projects generally receive correct funding which arrives on time throughout the year. A project usually lasts from 3 to 5 years. Approval of a project implies an institutional commitment to have funds available during that time.

It is difficult to distinguish between funds used for monitoring plans, programs, or projects. Table A.1 shows the human and financial resources which can be changed to monitoring and internal evaluation of all technical activities at INTA.

Table A.1 Estimate of monitoring and evaluation costs

	Number of professionals	% of time spent on M&E	Professionals/ years	Professional cost (@US$ 32,500/yr)
Central group	5	100	5.0	162.500
Assistants at Centers	18	25	4.5	146.250
Center Director	18	15	2.7	87.750
Exper. Station Director	50	25	12.5	406.250
Program Co-ordinator	15	30	5.0	162.500
Subprogram Coord.	30	30	10.0	320.500
Area Co-ordinator	120	10	12.0	390.000
Management staff (10% total)				164.575
Subtotal				1.843.325
Operational costs(20% salaries)				368.665
Total				2.212.090

Annual monitoring has generated multiple products at INTA. In the first place, the project and work plan reports are prepared for the use during the annual reviews. Review documents are written at each Center and these include proposals for the operational planning of future Center activities, such as reports on plans already executed, and on any changes in priorities and the allocation of economic resources for future exercises.

The conclusions drawn from monitoring and evaluating projects and work plans, which are prepared during the Internal Evaluations, constitute an important element decision making for those in charge and also the Center's Council.

The Center's policies and priorities are defined by taking into account the progress of activities programmed and products generated. At INTA, participating producers, i.e. the Institution's audience, participate in the Local Councils of the extension agencies and experimental stations, and in the Center's Councils. This implies a strong "social control" of activities completed and a continous analysis of the products obtained. The information is published and available both for internal political and management levels and for external levels. The results of monitoring and evaluation may be used to change or modify program preparation, and resource allocation, and in some cases, even to relocate staff, which projects are cancelled or plans abandoned.

INTA has institutionalized monitoring by means of a number of resolutions of the Governing Council; these have been accepted by all those who are involved in the institution. These dispositions are normally carried out. The responsibilities assigned to the different levels of authority are based on a decree of the national government, whereby Centers have the explicit responsibility for conducting monitoring. This approach was ratified by INTA's Governing Council in a resolution approved by the Institutional Evaluation Program.

Monitoring national programs and subprograms

National Programs and their subprograms, conduct two kinds of monitoring:

Monitoring activities and progress of the workplans to which they belong. Monitoring processes are organized by the program and subprogram coordinators, and implemented in the experimental units where they are situated. They often coincide with the Center's internal evaluations. These monitoring activities have led to decisions to end workplans that do not respond to program's needs, or to consolidate others that share similar objectives and do not justify continuation as separate workplans. In some cases, the need to undertake new activities through new workplans was analyzed.

Methods and techniques used in monitoring programs have been similar to those used for internal evaluations of the Regional Centers. These include meetings with the technicians and specialists working in the program, the discussion of progress, and a general analysis done by the co-ordinator of the program. Progress indicators are expressed at the level of activities and products.

Analysis and discussion of program objectives. Most programs and subprograms have organized meetings to discuss and monitor progress and revise objectives. These meetings are not very different from program planning activities, because the planning, monitoring, and evaluation process is done in simultaneous meetings in which the program is analyzed and discussed, and the new proposals enrich planning adjustments.

Given that national program activities are carried out through work plans prepared by the Centers, monitoring done by the Centers during internal evaluation cannot be totally separated from formal program monitoring. Therefore, monitoring resources, products, and use of results corresponds to the information already given for internal evaluations. National co-ordinators of programs and subprograms periodically visit and review program activities *in situ* in the different regions, producing analysis documents that are presented to technicians in charge, to the unit and Center directors, and to the NDDC&E.

The major difficulties and problems for monitoring programs have originated in the lack of precision in the objectives and goals of the proposal documents. As stated in the chapter on planning, these documents are being rewritten to improve accuracy level.

Evaluation

After the institutional change bringing about decentralization in 1986, an Institutional Evaluation System was organized to provide information to the Governing Council and the National Directorate on the progress of technical activities at INTA. The information was to be used by decision makers for formulating policies and institutional strategies, and for informing those in charge of the progress of their activities.

External evaluations

The institutional exposure of INTA to external influences, consolidated at policy level by the Center's Councils, was complemented by scientific and technological exposure through the Institutional Program of External Evaluations, following recommendations given by the Governing Council and external experts.

The evaluation covers all of INTA's structures: experiment units, research institutes, programs, subprograms, special units (such as projects for small farmsteads), technical cooperation, etc. From July 1990 until August 1992, 16 experiment stations were evaluated. Two more will be evaluated in October 1992. The objective of the external evaluations is to give the Governing Council elements to examine a critical analysis of the units' progress in relation to its operational plans and the degree of correspondence with broader institutional objectives and priorities. These evaluations are carried out by externally contracted evaluators.

Few national research institutions had no practical experience in control and evaluation that could serve as model for INTA. This forced INTA to develop its own proposal, adapted to the characteristics of an organization that does both research and extension, with different forms of regional operation, organization levels, and technical objectives for rural development.

The NDDC&E prepared a proposal to organize and develope an institutional evaluation program. It was analyzed and approved by the Governing Council, who delegated the executive responsibility to the NDDC&E.

The methods and techniques used for the external evaluations are based on the following mechanisms and instruments:

The operational plan. This serves as a general guideline and includes:

- Presentation of the annual evaluation program for the approval of the Governing Council.
- Selection of evaluators to make up the Evaluation Commission of the unit to be evaluated.
- Selection of documents to be given to the Evaluation Commission a month before beginning their mission.
- Implementation plan for the Evaluation Commission
- Evaluation report:
  - General characteristics of the evaluation report.
  - Procedure to be followed in preparing the report.

Framework. In order to provide a framework for the evaluation commission, general and specific terms of reference (TOR) are prepared, giving the Commission a series of priority themes. Each commission should these analyze starting with the general evaluation concepts that guide the program.

The main headings in the TOR are:

- The action program of the unit to be evaluated.
- The unit's relationship with the rural environment of its area.
- Achievements, progress and impact, in both research and extension.
- Scientific and administrative management at the different levels of responsibility.

Specific TOR are also provided with questions concerning the unit's organization, its responsibilities, activities, resources, and results.

External evaluation commissions are made up of an average of four or five external consultants for each experiment station evaluated. The evaluation commission is supported by two NDDC&E staff members, who make sure that the process follows the Governing Council mandate and the operational plan of the Institutional Evaluation Program.

The operational models of the external evaluations vary, but comply with a common general program for all the units. First of all, the commission receives information from the Center and the unit evaluated in presentations made by Center directors and by experimental station directors. The main projects of the unit are generally presented on the first day.

The evaluation commission is introduced to all researchers and extension agents in the unit, and then meets with them individually or in groups, without the directors.

During the week that the evaluation committee is working, it meets not only the professionals in the unit, but also the other professionals in the community, and producers, rural organizations, universities, and local government. Finally it meets the Center's Governing Council. The last two days are dedicated to preparing the report.

The cost of evaluating an experimental station is estimated to be approximately US$10,000.

The direct product of the external evaluation of each experiment station is the commission's report to the Board of Directors.

Once the report has been presented and considered by the Governing Council, it is sent to the Council and to the Center Director for analysis and so that proposals can be made on actions to be taken on the basis of the report's recommendations.

Changes resulting from external evaluations have been significant. In one case (Anguil Experiment Station), the conclusions of the external evaluation changed the whole operational plan of the unit and those responsible for carrying it out. In the other 16 cases, the external evaluation meant a review of all on-going activities. The proposal made in all cases was to intensify research-extension interaction, and to conduct microeconomic studies of the technological proposals. In the 1990-1991 evaluations, the external evaluation brought about regional program reviews of certain products; for example, potatoes in the Balcarce Experiment Station, and citric fruits and forestry in the Montecarlo Experimental Station.

In the case of the Paraná Experiment Station, the external evaluation resulted in a review by the Center's Council of priority objectives. In six experimental stations, external evaluations led to ratification of operational proposals for which the Center's Council had established priorities. Another result of the external evaluation was confirming that activities had been properly completed at the experiment stations of Mendoza-Yuto and Cerrillos in Salta, Manfredi in Córdoba, Mercedes in Corrientes, and Cerro Azul in Misiones. The evaluation at the Rafaela Experimental Station showed a lack of communication between researchers and extension agents, and its negative effect on their work.

Evaluation is now an institutionalized process at INTA. Approval in 1990 of the Institutional Evaluation Program by the Governing Council, as well as the annual evaluation programs, provided strong political support for the whole process.

No important problems have yet arisen, either in implementing the external evaluations or in using their results.

Final project evaluation

Final project evaluations have recently begun at INTA, to analyze results and products and to inform the Governing Council, which is responsible for approving and allocating resources. This evaluation was also intended to analyze the fulfilment of goals and objectives by those in charge of conducting activities, and by other participants.

The project evaluation analyzes what has been implemented and the results in relation to what was planned, the degree of participation of institutions and beneficiaries involved and the allocation and use of financial resources.

The NDDC&E is responsible for organizing final project evaluations. The Centers in whose experiment stations the projects operate are responsible for carrying out the evaluations.

In the specific case of projects for small producers, the responsibility for organizing the evaluation of those plans and projects was delegated by the NDDC&E to the Unit of Plans and Projects for Small Farms.

The participation of different levels of authorities for the evaluation takes place sequentially. At the first level, projects and results are evaluated by beneficiary producers. Participating institutions and project technicians constitute the second level. The Central Administrative Direction, experimental stations, and the National Direction participate at the third level. This participative methodology by level has been used to evaluate 12 projects for small farmstead producers.

In the cases of projects for small producers, the final project evaluations were carried out with technicians from the NDDC&E, the Centers, and

from the Unit of Plans and Projects for Small Farms. This means that the only additional costs over the basic operating costs were the technicians' travel costs, and their expenses during the days they stayed at the project zone during evaluation.

Reports have been prepared for each of the final project evaluations carried out. These reports are sent to the Center's Governing Council and also the NDDC&E.

Results of the final project evaluation have been used within the institution to analyze fulfilment of goals and objectives. This helped to plan new projects in the same areas, correct management errors and generate proposals for new projects. The evaluation has also been useful for informing the external organizations that helped fund the projects.

The final evaluation of all the projects approved by INTA's Governing Council has not been carried out. The final evaluation of projects for small producers has been completed for both approved and finished projects.

Technical auditing

Technical auditing is done in the different units of INTA where the National Direction judges that reasons or special circumstances justified an in-house, in-depth analysis of the unit's activities.

The objective of the audits is to find out directly the degree to which the operational units or projects are functioning well. The audits attempt to give an objective report of the situation to the National Direction. These audits are an important management function under the supervision of the National Directorate. They are *ad hoc* mechanisms, carried out whenever the National Directorate deems it necessary, and constitute a central evaluation mechanism in a highly decentralized organism.

The NDDC&E organizes and carries out the technical audits with the participation of the directors of the Assistant National Direction of Planning and Operations, the Program and Subprogram co-ordinators of National Environment technically related to the operational unit being audited, and the directors of units with similar profiles or related areas of interest.

To date, technical audits have been carried out at the Famaillá Experiment Station, at two Research Institutes (Soils and Microbiology, and Agricultural Zoology), and a Technological Cooperation Project. In all the cases, the audits produced an internal report, presented directly to the National Direction. As a consequence, changes in management personnel and in activities have taken place.

Appendix 7. Bibliography

References

ASHBY, J. 1992. Manual para la evaluación de tecnología con productores. Centro Internacional de Agricultura Tropical: Cali, Colombia.

BLALOCK, H. M. 1968. Theory building and causal inferences. **In**: Blalock, H.M. Jr.; Blalock, A. B. (eds.). Methodology in social research. McGraw-Hill: New York, USA.

BORGES-ANDRADE, J.E.; QUIRINO, T.R. 1989. Mensuração da qualidade de publicações e comparação com outras medidas de efetividade organizacional em pesquisa aplicada. Ciencia e Cultura. 41(2):138-145.

BORGES-ANDRADE, J.E.; HORTON, D. 1993. Planning, monitoring and evaluation in EMBRAPA, Brazil. Discussion Paper No. 93-10. International Service for National Agricultural Research: The Hague, The Netherlands.

BRIONES, G. 1982. Métodos y técnicas de investigación para las ciencias sociales. Trillas: México D.F., México.

DUVERGER, M. 1981. Métodos de las ciencias sociales. Barcelona, España.

EMBRAPA. 1990. Projeto: Diagnóstico global da EMBRAPA. Empresa Brasileira de Pesquisa Agropecuária, Secretaria de Administração Estratégica. Brasilia, Brasil.

FONSECA, S. 1990. La evaluación en instituciones de investigación agrícola. Centro de Desarrollo de las Investigaciones de la FAO. Documento de Trabajo No. 8. Food and Agriculture Organization of the United Nations: Rome, Italy.

HORTON, D.; PETERSON, W.; BALLANTYNE, P. 1993. M&E Principles and Concepts. **In:** Horton, D.; Ballantyne, P.; Peterson, W.; Uribe, B.; Gapasin, D.; Sheridan, K. 1993. Monitoring and evaluating agricultural research: A sourcebook. CAB International in association with ISNAR: Wallingford. U.K.

de evaluación de actividades técnicas. Instituto Nacional de Tecnología Agropecuaria: Buenos Aires, Argentina.

McLEAN, D. 1988. Monitoring and evaluation in the management of agricultural research. Working paper No. 14. International Service of National Agricultural Research. The Hague, The Netherlands.

MARTINEZ, R.; HOLM, K.; NUBNER, P. 1983. Introducción a los métodos de la sociología empírica. Alianza: Madrid, España.

MURPHY, J. 1993. Monitoring and evaluation in agricultural research: Concepts, organization and methods. Informal Report. International Service for National Agricultural Research: The Hague, The Netherlands.

NOVOA B., A.R. 1989. Agricultura, tecnología y desarrollo: Cómo se evalúa la investigación agraria en América Latina. Tercer Mundo Editores en asociación con CIID: Santafé de Bogotá, Colombia.

NOVOA, A.R.; HORTON, D. 1994. Planificación, seguimiento y evaluación de la investigación agropecuaria en las Américas: Una síntesis. **In**: Novoa B., A.R.; Horton, D. (eds.). Administración de la investigación agropecuaria: Experiencias en las Américas. Tercer Mundo Editores en asociación con ISNAR y PROCADI: Santafé de Bogotá, Colombia.

PARDINAS, F. 1977. Metodología y técnicas de investigación en ciencias sociales. Siglo XXI: México D.F., México.

POSADA, R. 1994. Planificación, seguimiento y evaluación en ICA, Colombia. **In**: Novoa B., A.R.; Horton, D. (eds.). Administración de la investigación agropecuaria: Experiencias en las Américas. Tercer Mundo Editores en asociación con ISNAR y PROCADI: Santafé de Bogotá, Colombia.

QUIRINO, T.R.; CRUZ, E. R. DA; SOUZA, G. DA SILVA. 1992. O processo de produção de conhecimento em organizações de pesquisa agropecuária: Diagnóstico de projetos. (Versión preliminar). Empresa Brasileira de Pesquisa Agropecuária, Secretaria de Administração Estratégica: Brasilia, Brasil.

QUIRINO, T.R.; COQUEIRO, E.P. 1985. Um exercício de avaliação da qualidade da pesquisa agropecuária na EMBRAPA. Revista de Economia Rural 23(3): 351-368.

ROMANO, L. 1987. Economic Evaluation of the Colombian Agricultural Research System. Ph.D. Thesis. Oklahoma State University: Stillwater, USA.

ROMANO, L. 1988. Evaluación económica de la investigación agropecuaria. Teoría y práctica. Boletín Técnico No. 176. Instituto Colombiano Agropecuario: Santafé de Bogotá, Colombia.

ROMANO, L. 1994. Planificación, seguimiento y evaluación en el ICTA, Guatemala. **In**: Novoa B., A.R.; Horton, D. (eds.). Administración de la investigación agropecuaria: Experiencias en las Américas. Tercer Mundo Editores en asociación con ISNAR y PROCADI: Santafé de Bogotá, Colombia.

ROSSI, P. H.; FREEMAN, H. E. 1985. Evaluation: A systematic approach. Sage: Beverly Hills, USA.

.URIBE, B.; HORTON, D. (eds.). 1993. Planeación, Seguimiento y Evaluación de la Investigación Agropecuaria: Experiencias en las Américas. Informe del Taller Regional. CIMMYT, México. 15 - 22 Octubre, 1992. Servicio Internacional para la Investigación Agrícola Nacional. The Hague, The Netherlands.

Literature consulted

COOMBS, C.; DAWES, R.; TVERSKY, A. 1970. Mathematical Psychology. Prentice-Hall: New Jersey, USA.

HOGG, D.R. 1994. Planificación, seguimiento y evaluación en INTA, Argentina. **In**: Novoa B., A.R.; Horton, D. (eds.). Administración de la investigación agropecuaria: Experiencias en las Américas. Tercer Mundo Editores en asociación con ISNAR y PROCADI: Santafé de Bogotá, Colombia.

ISAAC, S.; MICHAEL, W.B. 1974. Handbook in research and evaluation. Edits Publishers: San Diego, USA

LAZARSFELD, P.F.; MENZEL, H. 1969. Sobre la relación entre propiedades individuales y colectivas, **In**: Korn, F. et al. Conceptos y variables en la investigación social. Nueva Visión: Buenos Aires, Argentina.

ZAPATA, V. 1995. Training Trainers in Agricultural Research Management. International Center for Tropical Agriculture (CIAT), in association with ISNAR. Cali, Colombia.