RESEARCH EVALUATION REQUIREMENTS IN THE CGIAR

April 1995

Julian M. Alston and Philip G. Pardey

Department of Agricultural Economics, University of California at Davis, and International Food Policy Research Institute, Washington, D.C., respectively.

A Report Commissioned by the Impact Assessment Task Force
for the Public Awareness and Resource Mobilization Committee (PARC)
of the
Consultative Group on International Agricultural Research (CGIAR).

This report was prepared under severe time constraints, aiming to promote a free discussion about the issues raised, and thereby to lead to the development of some clearer directions for change. We are grateful for the comments and assistance we received from the members of the Impact Assessment Task Force and colleagues throughout the system. Notwithstanding this help, the views expressed in this report are preliminary, representing the personal views of the authors only. They are not attributable to any others at the University of California, IFPRI, PARC, or the Impact Assessment Task Force.

(revised April 25, 1995)

Table of Contents

		Page
1.	Introduction	1
2.	Conceptual Overview	3
3.	Prior Practice	12
4.	Principles for Designing Evaluation Institutions	18
5.	Current Capacity	20
6.	Policy Options	28
7.	Reference	32
8.	Appendix	34

1. Introduction

Like agricultural research institutions worldwide, the Consultative Group on International Agricultural Research (CGIAR) is being asked to do more with less. Research evaluation, sometimes called impact assessment, may help in two ways:

- *by improving the information on which donors make decisions about the amount of resources to allocate to the system, or parts of it, thereby affecting the total resources available for research*
- *by providing information that improves the allocation of research resources provided to the system, thereby increasing the benefits from a given amount of research resources.*

Hence, in October 1994, as part of its reform package, the CGIAR proposed to strengthen its research evaluation activities. This proposal was endorsed at the special Ministerial-Level Meeting held at Lucerne, Switzerland in February 1995, with a specific recommendation to "strengthen the assessment of its performance and impact by establishing an independent evaluation function reporting to the CGIAR as a whole."[1]

In parallel with these developments, the Public Awareness and Resource Mobilization Committee (PARC) of the CGIAR established an Impact Assessment Task Force to:

- assess what centers are currently doing in impact assessment, the methodology they follow, and the data available
- determine what impact should be assessed (including to serve as an instrument of accountability) and at what level (international research, CGIAR, centers)
- bring together expertise from within (including TAC) and outside the CGIAR to develop a systematic and system-wide methodology and process of impact assessment
- link impact assessment with public awareness and resource mobilization
- propose a mechanism for systematic impact assessment.

Terms of Reference

This report was commissioned by PARC through the Impact Assessment Task Force in the context of these broader developments. The specific elements to be addressed are:

- identify the needs for and strategic objectives of impact assessment (defined here as ex ante and ex post) within the CGIAR at the center and system levels, with due regard for the mission statement and the range of activities across the centers
- in the light of these needs and objectives, outline the range of approaches useful to impact assessment in the CGIAR
- identify the current roles and existing efforts of CGIAR entities in impact assessment

[1] This section, and some other parts of this report, draw on Özgediz (1995).

- highlight current and future needs and objectives not now addressed at the CGIAR level, suggest approaches and methods of assessment useful to deal with these and the indicative amounts of resources this may require
- based on the above, lay out the key principles and some options for institutionalizing impact assessment within the CGIAR.

Objectives and Scope of this Report

The CGIAR is involved in a number of activities, including (a) various types of agricultural R&D (ranging from more basic research, through applied and adaptive research, to outreach and extension activities),[2] (b) strengthening national agricultural research systems (NARS) in less-developed countries (LDCs) through training of scientists, (c) strengthening institutional development in developing-country NARS, (d) sponsoring research networks among NARS, and (e) developing and maintaining global stocks of agricultural genetic resources.

Evaluation methodologies and data bases are relatively well developed for certain types of agricultural R&D, and virtually nonexistent for most of the other activities. While the fundamental principles are the same for evaluating all CGIAR functions, the specific information requirements, methodological details, and practical problems differ. Given the very short time-frame available, in preparing this report we have focused on evaluation of research, the primary function of the CGIAR (accounting for about one half of the system's core expenditures).[3] Since research is the most important area of funding, clearly it is the most important area to get right and a logical place to begin impact assessment.[4] We recognize, however, that many seek a comprehensive and comparable assessment of all CGIAR activities so, where we can, we introduce more specific ideas and arguments related to evaluating the non-research functions.

Our ultimate objective is to set out some options for developing new institutional arrangements for research evaluation within the CGIAR system. In considering CGIAR options for institutionalizing the evaluation of research, we have in mind the twin fundamental purposes for such evaluation: (a) justifying funding support, and (b) improving resource allocation decisions. We first provide a brief conceptual overview of research evaluation issues and

---

[2] For research evaluation, one can adopt alternative characterizations of the CGIAR's research by identifying either its thematic orientation (e.g., genetic improvement, crop and animal protection and management, agricultural policy, or environmental consequences of agriculture) or where it has its effect in the agricultural input-production-postharvest continuum.

[3] The remaining resources are deployed as follows; 25% for research management, 10% for research support, and 16% for institution building.

[4] As well as being the primary area of interest, agricultural research evaluation is also the area where knowledge in general is most well developed and where we have the greatest relevant expertise. In addition to being able to draw on an extensive professional literature in the area, we have worked in practical research evaluation and priority-setting roles. Comparable experience and resources are not available for evaluating the economic impact of the institutional development and coordinating roles of the CGIAR.

methods in section 2, in relation to the strategic role of research evaluation within the CGIAR. We discuss the different types of research evaluation activities that might serve the various demands for accountability, to shore up funding, or to support decision making at various levels within the CGIAR, given its current (or potential) institutional setting and structure, and in the light of its mission and objectives. In section 3, we discuss the prior practice for research evaluation within the CGIAR, and identify areas of weakness. In section 4, we discuss alternative approaches that might be adopted, in principle, to develop an institutionalized process of research evaluation and priority setting within the CGIAR. In section 5 we identify current capacity within the CGIAR system for conducting evaluation and priority-setting studies, including human capital, data, and methods. In section 6 we describe a set of feasible options and discuss their main strengths and weaknesses. Finally, section 7 concludes the report.

2. Conceptual Overview

In designing an institutional arrangement for evaluating agricultural research, it is necessary first to consider what the purposes of such evaluation may be (i.e., how the results would be used), and within the CGIAR system there are several purposes calling for different types of evaluation. Evaluation involves assessing performance against objectives. Appropriate objectives for research must be defined before proceeding to the design of evaluation mechanisms. This is so whether the intent is to evaluate the effects of past research (ex post evaluation) or the effects current and future research (ex ante evaluation). The objectives define the set of indicators by which the impacts of CG research are to be assessed. We review the implications of the CGIAR mission statement and donor requirements, against some economic arguments, for the definition of objectives for the CGIAR in research to be used as a basis for impact assessment.

Purposes for Evaluation

The two primary purposes of research evaluation from the point of view of participants in the CGIAR are (a) to obtain and justify funding support and (b) to support making decisions about the allocation of funds among alternatives. From an external (e.g., donor) perspective, the value of research evaluation may be to inform decisions about the appropriate amount of overall support for the system and decisions regarding the types of CGIAR activities to support in future. Whether it is backwards looking (ex post) or forward looking (ex ante) in nature, to have value itself, research evaluation work must be useful in some decision that affects future outcomes. It is difficult to justify the expenditure of resources on ex post evaluation work that is not intended to serve some decision or other.

Types of Decisions Served by Evaluation[5]

Different types of evaluation relate to different types of decisions. The highest level of research resource decision making is at the level of donor support for the entire system or for particular centers. The information requirements of donors may be considered an important determinant of the types of evaluation to undertake at this level. Such requirements include information for public awareness which extends beyond the immediate donor community.

Donor Decisions: Often, very broad indicators have been used. Ex post studies provide indicators of aggregate rates of return or benefit-cost ratios (or even simply rates of uptake of new technologies or of agricultural productivity growth) for particular programs of research, individual centers, or the system as a whole. Such indicators are often taken as evidence that past investments have been effective and, by implication, that future investments will be effective as well. Alternatively, ex ante projections of population growth relative to the capacity of the food system, or of the rate of loss of natural resource endowments, have been used to point to a "need" for maintaining or increasing research efforts in order to meet future demands (e.g., McCalla 1995; various papers produced in IFPRI's initiative on "A 2020 Vision for Food, Agriculture, and the Environment").

Information of these types might be effective in justifying past support and mobilizing some funding for the future. But, increasingly, those who finance research investments are asking for more specific information about the payoff. And, more specific decisions (than whether to fund research or not) require more specific, detailed, information about the consequences. For example, Australia's choice about whether to provide an additional $1 million in funding for either CIMMYT or IRRI (or neither of them) may require specific information about the past or likely future consequences of CIMMYT and IRRI research for productivity of wheat and rice in a particular group of developing countries (e.g., in Asia and the Pacific) as well as in Australia itself.[6] Such interest of donors in the details may extend as far as specific programs of research.

CGIAR Choices: Similarly, within the CGIAR system, decisions must be made about which centers to support and how resources should be allocated among them (indeed, how many centers and how they should be structured and relate to one another). Some decisions relate to programs of research that involve parts of the work programs of multiple centers (such as STRIPE reviews); others involve the entire work program of one or more centers. Such decisions require information on the expected future performance of the alternatives in relation

---

[5] We focus on the two fundamental decisions to be served by evaluation, and this leads us to focus this paper on the quantitative approaches to evaluation, based on economic principles. Other approaches are better suited to other types of decisions, such as day-to-day management of scientists. They may serve a useful monitoring role or provide data that can be used in a quantitative assessment, but they cannot substitute for the methods discussed here. Some of these other approaches are discussed by Horton et al. (1993).

[6] See ACIAR (1994).

to the mission of the system—either ex post assessments of past performance or ex ante assessments.

Within centers, similar types of distinctions arise. Some decisions involve complex comparisons in a comprehensive assessment while others involve more narrowly defined alternatives. One type of decision, for example, is the allocation of discretionary resources among programs within a center, calling for an assessment of the relative merits of the center's entire portfolio. Another type of decision is which projects to undertake within, say, a rice breeding program. For such detailed decisions, structured scientific peer review may be the best approach, whereas for the programmatic decisions, objective economic assessment is often more applicable (e.g., see Alston, Norton and Pardey 1995).

Approaches to Evaluation

In this context, one issue is how to make best use of the scientific and economic information that exists in the system being managed. Making centralized (top down), research decisions based on formal assessment presumes that managers can obtain the relevant information at reasonable cost. An alternative is a decentralized approach for some decisions, competitive approaches for others (requiring scientists, or the institutions in which they work, to compete for funds in terms of a defined criterion or objective), and centralized decisions for others. Where decision-making responsibility is devolved, a different type of evaluation may be required in order to hold subordinate research managers accountable.

As well as differing according to the type of decision being made, and the objectives of who is making the decision, research evaluation differs according to the type of research being undertaken, because the types of information that are available differ. For example, assessing the benefits of basic or pretechnology research is intrinsically more difficult than assessing the benefits of applied or adaptive research. Similarly, assessing the benefits of social science research, or other research that leads to disembodied technical change, is harder than assessing the benefits from research whose results are embodied into new machines or genetic material. Even when the criterion for research evaluation is the same for all programs (e.g., net economic benefits), quantitative benefit-cost analysis may be infeasible for some types of research activity and choices must be made about how to evaluate and compare alternatives when some benefits can be quantified and some cannot.

Research programs can either be continuing or discrete. Research evaluation itself is a type of research program and it, too, can be ongoing or a one-shot exercise. One-shot, discrete, evaluations can be valuable for informing the specific decisions that they are designed to serve. Often they are used as a basis for other types of decisions (we often rely on past evaluations) for which, over time, they become increasingly inappropriate. Comprehensive, system-wide reviews tend to have this quality. Continuing assessments are more likely to provide information that is contemporaneously useful for the continuing decisions about which projects and programs to emphasize.

A research evaluation institution could be developed to undertake the entire range of evaluation activities, ranging from system-wide ex post assessments to program reviews or to explicit ex ante assessments designed to support intra-program resource allocation decisions. Alternatively, one could set out to develop a culture in which research evaluation is carried out primarily by those who are making the decisions (different people for different decisions at different levels in the system). In the latter approach, a central capacity might be developed to support central (resource allocation) decisions, to bring about the development of standardized methods and criteria, to manage or at least mobilize centralized data bases, and to advise decentralized decision makers on methods and procedures.

Evaluation Criteria: The CGIAR Mission

Since decisions involve choosing among alternatives, the information about the consequences of alternative research decisions must be put on a comparable footing. This means research alternatives must be assessed in relation to a small number of criteria that pertain to the objectives for the research. The question of appropriate objectives for a research system (and the CGIAR in particular), and how to measure performance against those objectives as a basis for decisions, has been controversial.

The current mission and goals statement of the CGIAR, adopted in October 1990, reads:

> *Through international research and related activities, and in partnership with national research systems, to contribute to sustainable improvements in the productivity of agriculture, forestry and fisheries in developing countries in ways that enhance nutrition and well-being, especially of low-income people.*[7]

The intent of this statement is to define the *broad* goals of the system. But this level of generality limits the usefulness of this statement as a basis for designing meaningful research evaluation strategies and measures of impact against which to judge the performance of the system. And, the multiple objectives reflected in this statement raise additional real difficulties in this regard. Thus the CGIAR goals, as described above, require considerable elaboration before meaningful indicators of impact can be devised.[8]

Does the CGIAR mission statement mean, for instance, that research programs that are very effective in enhancing productivity should not be undertaken if they do not also enhance nutrition? What if a particular development would not be sustainable, but would be very

---

[7] The CGIAR mission statement has changed over the years. This version of the statement explicitly broadened the commodity coverage beyond the food-crops orientation of previous versions to include agriculture (i.e., crops and livestock), forestry, and fisheries while retaining the "sustainabilty" objective that was incorporated into the 1986 statement.

[8] For TAC's interpretation of the implications of the current mission statement for research evaluation and priority setting purposes see TAC (September 1990, chapter 7).

effective in alleviating poverty for, say, the next 10 years; would it be desirable? The problem is that real-world alternatives might not contribute to *all* of the stated goals and some might contribute to one at the detriment of the other. Hence, tradeoffs are involved; but the dimensions of the tradeoff are difficult to quantify both on the "supply side" (how much productivity *must we give up* in order to achieve an income distribution objective?) and the "demand side" (how much productivity are we *willing to give up* in order to achieve an income distribution objective?). Adding more objectives exacerbates the problem.

Reasonable Objectives for Research

In considering the impacts of any CGIAR activities, including agricultural R&D, we could consider a range of objectives (or types of impacts, or criteria) to be considered in making an assessment. Often stated objectives for research include such things as sustainable agricultural productivity growth, nutritional status of the poor, food security, or employment (e.g., as in the CGIAR mission statement above). Many economists would argue for translating all relevant effects of research into a single criterion, the overall net benefits from research. The use of a single criterion has a number of important advantages, including that it simplifies the problem of evaluating the consequences of research and comparing alternatives, and that thereby it reduces the costs of decision making.

Economic Efficiency: The *economic* justification for government intervention of any type in agricultural R&D is that economic efficiency will improve as a result. In this context economic efficiency is an inclusive concept that refers to the achievement of the greatest net benefits for the society as a whole, taking a broad view of net benefits.[9] In particular it means not seeking to give benefits to a particular group at the expense of society as a whole (of course, avoiding preferential weight is not the same as giving zero weight, it means giving equal weight). Applying the same principles and arguments also leads to the conclusion that the appropriate objective of public-sector agricultural research is the pursuit of economic efficiency, the greatest net benefits overall. This is not to say that other objectives are illegitimate, irrelevant, unimportant, or ignored. Rather, there are likely to be alternative, less costly ways to achieve these other objectives than by moving the agricultural R&D portfolio away from programs that will maximize total net benefits (broadly defined).

Agricultural research tends to be a blunt and ineffective instrument for most non-efficiency objectives (such as food security, income distribution, the balance of trade, and, perhaps, objectives related to non-economic efficiency aspects of the environmental consequences

[9] For instance, it includes benefits such as sustainability (to the extent that there are net benefits from the development and adoption of more sustainable resource-use patterns), environmental objectives (e.g., where R&D can lead, economically, to a reduction in pollution), and nutritional objectives (to the extent that there are net benefits from R&D directed towards improving dietary quality and health). Such nonpecuniary benefits are included along with pecuniary benefits, in a broad concept of net national benefits. The idea of economic efficiency is to make those net national benefits as great as possible. For more details on these arguments, see Alston and Pardey (1995) and Alston, Norton and Pardey (1995).

of agriculture). Hence, in general, we would argue against the use of agricultural research as a policy aiming to achieve a particular effect on income distribution, for example, both because the income distribution effects of research are very difficult to anticipate or control and because there are more effective policies available for changing the distribution of income. On the other hand, in a context where other policies are not available or are ineffective (e.g., such as in relation to poverty and malnutrition in Africa), agricultural research may be the best way of attempting to address a social problem. The same (or a similar) research portfolio might be chosen if the criterion were simply maximizing net research benefits in a context where alleviating malnutrition is regarded as an important benefit.

As a practical matter, the CGIAR's allocation decision primarily involves a choice of the commodities to research, the agroecologies to target, and the types of technologies to produce (considering the balance between improving agricultural productivity and addressing the pollution and degradation consequences of agriculture). The critical question is the extent to which the research portfolio ought to be deliberately distorted away from the program that would maximize economic efficiency (i.e., the greatest net benefit to all) in order to achieve some other objective (e.g., the greatest benefit to a particular group).[10]

One important source of confusion concerns how to decide what things ought to be included and what ought to be left out of the benefits calculus. The correct answer is that all effects that we care about, and that are truly attributable to the research, ought to be included. Some effects, however, ought to be left out if we believe, for instance, that other policies can (or perhaps it is sufficient to say, could) be used to balance out the consequences of research. In such cases, the "side-effects" are not so much attributable to the research but to other policies (or to a failure to implement other policies). For instance, it could be argued that a certain new technology may involve increased costs of environmental pollution that ought to be charged against the other research benefits when the technology is adopted. On the other hand, if environmental policies were chosen appropriately, those additional costs may not arise.

Sustainability: The issue of "sustainability" research is a good case in point. There is no doubt that agricultural research can have environmental consequences and that, in the past, some of those consequences have been treated inappropriately or ignored (Alston, Anderson and Pardey 1994). Appropriate accounting for the value of natural resource stocks and so on is necessary, even when the analysis is concerned only with the economic consequences of research. Some would say that, in addition to the economic benefits thus identified, environmental impacts should get extra weight and research should be pushed in the direction of sustainability (however poorly defined) since farmers (or others) do not have appropriate

[10] In those countries where poverty is pervasive, maximizing the overall benefits from R&D is likely to be the best way of tackling the poverty problem as well. Many seem to think that pursuing an objective of maximizing the total social benefit from R&D ipso facto means that poverty problems are somehow being relegated to a lower level of concern (or ignored altogether). Such is not the case.

incentives to preserve resources.[11] As argued by Alston and Pardey (1995), however, the best policy to correct such problems may be to modify the incentives. Biasing the research portfolio in the direction of environmentally friendly projects will be ineffective unless the farmers have incentives to adopt the results of such research. The first-best policy (correcting incentives) may be necessary to make the research policy effective; but the first-best policy may, at the same time, make the second-best correction (i.e., deliberately distorting the research program away from the one with the greatest payoff) unnecessary. A policy that forces farmers to pay for the pollution they create should (and demonstrably does) put pressure on scientists to develop technologies that reduce these social pollution costs along with other costs of production. Thus, the invisible hand of the market can be encouraged to direct the efforts of agricultural scientists dependent on public and even private (e.g., industry) sources of funds.[12]

CGIAR Constraints: The arguments supporting the use of a single economic-efficiency criterion for evaluating agricultural research have not been widely accepted in previous research evaluation and priority-setting work in the CGIAR, even by some economists. Partly this reflects a confusion between the requirements of donors for funds to be directed toward a particular center, or even a particular program of research, *a constraint*, and what criterion ought to be used for allocating resources that are more within the CGIAR's control, *the objective*.

Some of this confusion comes from the *consultative* structure and decision-making processes that exist at the apex of the organization. This "participatory" mode of operation keeps the system more directly engaged with its donors. A disadvantage is that donors may therefore be able to exercise too much influence over the detailed research resource allocation decisions, including some for which they do not have the relevant information or expertise. The same concerns may arise through the increased attention on systemwide accountability processes which provides another avenue for donor influence. Of course, donors should have *some* influence, but they are not the only clients and it cannot be assumed that interests of a specific donor necessarily coincide with systemwide objectives.

Thus, the CGIAR operates subject to a number of constraints. If the CGIAR were to aim simply to maximize global net benefits from research, without regard to where in the world the benefits were received, it might give all the budget to a particular center, say IRRI, an outcome that may engender very little benefits in Africa (although there would be some benefits in Africa). Part of the CGIAR challenge is that its donors have specific objectives in relation to

---

[11] Sustainability is a buzzword: it means different things to different people. Tomasi (1991) provides a good discussion of the use of this terminology in the context of R&D policy. We prefer using more explicit terms to connote particular relationships between technology, production, and natural resource stocks.

[12] A significant problem is that we do not adequately understand the nature and size of these environmental consequences. Research aimed at establishing these relationships (as distinct from fixing the problem) is a candidate for public support, but here, too, there is an allocation problem regarding the amounts and types of research to conduct. The basic research resource allocation principles can be used to think through this problem as well (Alston and Pardey 1995 discuss this issue).

the geopolitical relevance of and distribution of benefits from its research.[13] Subject to the constraint of donor's requirements, however, the right criterion is to allocate the research resources to achieve the greatest net benefits, giving appropriate regard to the availability (or not) of other effective policies for dealing with other concerns. In time, it may be possible to inform the donors about the nature of the trade-offs, in terms of foregone net income, involved in directing research toward objectives other than total net benefits. In this way, the politically motivated constraints that donors place on the system may themselves be subject to change.

Criticisms of Evaluation

Common criticisms of quantitative evaluation of research benefits, either ex ante or ex post, are that it is:

- too demanding and difficult
- too data intensive
- not well understood by the decision makers (i.e., a black box)
- incomplete in that important effects cannot be quantified.

These criticisms are valid. To be worthwhile, formal research evaluation does demand a lot of effort of skilled analysts, and certain types of effects are difficult to quantify. The formal methods are applicable only to a subset of the questions that arise in the complex that constitutes the CGIAR. As Alston, Norton and Pardey (1995) have argued at length, however, it is worthwhile to develop a culture of an "economic way of thinking" about research resource allocation, even if a formal optimization is precluded by incomplete data. Even if the benefits cannot be quantified accurately or easily, it should be helpful to have the purpose of maximizing benefits in mind when considering alternatives. Quantifying some of the benefits helps. And, in the CGIAR system where the work is relatively applied, a relatively high proportion of benefits are, at least in principle, quantifiable.

New Software: In relation to the charge that evaluation is too demanding and difficult to be worthwhile, it is worth noting that the costs are changing. In recent years, a number of programs and software packages have been developed for evaluating returns to research. It is now possible for people with little formal training in economics to become adept in computing rates of return to research using a minimum of economic and scientific data. In addition, the methods have been well documented (e.g., Alston, Norton and Pardey 1995). While these methods have improved greatly, and become more accessible, it is still important to have

[13] Enlightened self interest is an implicit, if not explicit, feature of the involvement of many countries in the international agricultural research system, and national interests in research aid imply some restrictions on where and how the aid is spent. For instance, the Jackson Report recommended that the Australian development co-operation program be focused sharply on the countries of the Asia and Pacific region that are important to Australia politically and economically. Economists and others have recently begun to pay greater attention to the issue of "doing well by doing good" as an element of donor motivation in agricultural research (e.g., Ryan 1987; Tribe 1991, 1994).

professional economists involved in appropriate ways. Some of the difficult, technical, economic questions that arise in research evaluation cannot be resolved as a matter of course in a menu for non-economists. That formal evaluation is not *too* demanding can be seen in Australia where, since it has been introduced as a condition for funding, a large proportion of scientists have found becoming adept at it worthwhile.

Supporting versus Making Decisions: Some decision makers balk at the apparent complexity associated with "sophisticated" evaluation methods, thus giving rise to the "black-box" concern noted above. This concern stems in part from a failure to distinguish clearly between those activities and persons actually carrying out the analyses (i.e., the decision support structure) and those using the information for allocating research resources (i.e., the decision-making structure). The trick is to achieve the optimal division of labor and responsibilities involved in the decision-support, decision-making continuum, but to do this in a way that truly engages all of the participants in the process in an informed and active dialogue.

For this to work, analysts must do relevant analyses and present the information to decision makers in a timely and comprehensible way. Similarly, decision makers need to understand the rudimentary economic principles involved in evaluating research and thinking through the tradeoffs involved in making research priorities and allocating resources. There must be an active and on-going attempt to have a meeting of the minds for more quantitative decision-making processes to be worth the effort.[14]

Multiple Objectives

We have suggested that if the objectives for research can be reduced to a single criterion, the total net benefits to all (i.e., economic efficiency), then the evaluation problem and decision making will be much simplified. Some object to this in principle. Others who fundamentally agree *in principle* with adopting a single objective for research, reject it in practice on the grounds that it is politically unacceptable. Unless those who pay for the research can be convinced to abandon multiple, competing (and sometimes incompatible) objectives, it may be judged infeasible (or unwise) to adopt the economic efficiency criterion even for decisions that take effect within the CGIAR system, beyond the direct attention of donors.

The alternative—multiple objectives—makes life more difficult both for the analyst employed to evaluate research and the decision makers who would use the results. To conduct an economic evaluation of research, ex ante or ex post, requires documenting considerable detail on the scientific outcome and its consequences for variables such as yield, production, availability and price of food, and so on. This information is intrinsically interesting and can

[14] It does not follow that research managers or senior decision makers necessarily need to know all the details involved in doing the quantitative assessments, but they should understand the basic economic principles involved and it would help to have a clear notion of some of the related jargon. In an analogous fashion one would expect that for them to be effective in managing or allocating resources to a breeding program the managers should appreciate the basic steps involved in breeding and have a grasp of some of the terminology used by breeders.

be made available as a matter of course as a product of the evaluation activity (especially in consideration of the fact that different audiences appreciate and respond to different types of information). But other objectives may require other information that is much harder to assess. For instance, to establish the likely consequences of a new rice variety for nutritional status of the urban poor, or for variability of farm incomes (or the chances of famine), requires much more information, and is very much more difficult to assess than the net benefits because so many other uncontrolled factors become relevant. The objections that research evaluation is difficult, demanding, and data intensive are much more relevant when contributions to multiple social objectives are being considered.

Multiple objectives mean that it is necessary to quantify tradeoffs. This is not always possible in practice; it is almost always difficult. For instance, some have suggested that the way to help poor Africans in marginal areas is to conduct research into agriculture in marginal production areas. The opportunity cost of research into cropping technology for the arid zone is the potential benefits from research into something else. The past record suggests that it is difficult indeed to make much progress in arid-zone cropping. One might well conclude that even though some gains are possible in the arid zone, the resources should be directed to other areas (e.g., irrigated or higher potential rainfed areas) where much greater income gains are likely (perhaps in the same countries).

Some would argue that tradeoffs should not be involved; we should undertake all research with a net social benefit (however measured). The reality is that research resources inevitably are scarce, and usually the right (second-best) criterion becomes equating rates of return among alternatives at a rate much greater than the social opportunity cost of funds (if they were unconstrained). More of everything is not an option; a choice must be made.

Absent evaluation, such tradeoffs cannot be quantified. With evaluation we can determine how much national (or global) income must be foregone (by forgoing more productive research options) in order to achieve an income gain to a particular group. Previous experience has taught us that when presented with such information, decision makers are not prepared to give up too much net income to benefit a narrow minority.

3. Prior Practice

The present round of efforts to institutionalize research evaluation efforts within the CGIAR is not starting from scratch. Özgediz (1995) details the substantial amount of work that has been done at the system level to evaluate CGIAR research and to provide information to TAC, donors, and to others involved in setting priorities for R&D.

System-Wide Studies

The CG Secretariat commissioned two system reviews, the last one completed in 1986, as well as an "impact study" at the system level. The impact study was carried out under the supervision of an external project manager and a team of consultants during the period 1984-

1986. It generated 24 CGIAR Study Papers that described the nature of CGIAR research and its various "effects," mainly on a country-by-country basis and occasionally on thematic lines. The study also produced a synthesis report in 1987 and several summary reports in 1988.[15] While providing a useful and reasonably comprehensive historical and contemporary account of CGIAR activities, these studies were largely descriptive in nature. They stopped short of placing an economic value on the activities being evaluated.

The CGIAR's Technical Advisory Committee (TAC) has sponsored a number of initiatives to review the whole portfolio of CGIAR activities and establish strategic priorities and resource allocation guidelines for the system. These efforts have increasingly sought to augment informed judgement with more formal impact assessment and quantitative priority-setting methods. The most recent round of priority-setting deliberations was led by TAC's Standing Committee on Priorities and Strategies. It engaged a number of consultants based in Rome and developed and applied a *scoring* method to assist in its deliberations. The ecoregions currently employed by TAC were conceived as a part of that effort.[16]

TAC's Scoring Model: The scoring approach to priority setting developed by TAC is described in some detail in Gryseels et al. (1992) (see, also, McCalla and Ryan 1992). It begins by developing a "baseline" allocation of research resources across ecoregions, based on its congruency with the estimated value of production, number of poor, and area of "usable" land in each ecoregion (this congruency was relative to a composite index formed by summing across the three criteria, each of which was given a weight of one third). A set of modifiers were applied to the baseline priorities to arrive at a final set of priorities by region and commodity.[17] The final set of priorities were used to make resource allocation recommendations regarding the balance of effort and research resources across regions, commodities, and, ultimately, centers.

Even the best, most carefully conducted, scoring study is subject to some general problems as discussed by Alston, Norton and Pardey (1995, pp. 463-500), which mean that the results are highly subjective and difficult to replicate. As a consequence, the main value of a particular scoring study is at the point in time and context in which it takes place; those engaged in a scoring study probably do learn about the alternatives. But such studies add little to the stock of institutional knowledge or expertise that can be built on in subsequent studies. One of the fundamental limitations of the TAC approach is that it makes no attempt to incorporate meaningful ex ante evaluation procedures into its priority-setting deliberations.

---

[15] See Anderson, Herdt and Scobie (1988) for a listing of the documents produced by the CGIAR Impact Study.

[16] See Wood and Pardey (1995) for an evaluation of the ecoregional concept as it has been implemented by TAC and others.

[17] These modifiers were to take account of "(1) the special nature of the CGIAR as an international organization; (2) alternative sources of research supply; (3) the strength of national research programmes—capacity of the national system and country size; (4) the nature of self-reliance; (5) concerns for the efficiency of research; (6) equity issues—intensity of malnutrition and GDP per caput; sustainability—urgency, magnitude of deforestation and soil degradation risk; and (8) special issues." (Gryseels et al. 1992, p. 75)

Sub System-Level Evaluation

A substantial amount of research evaluation work has been (and continues to be) done within individual CGIAR centers. Some of this work involves ex post research evaluation studies of specific technologies or (commodity) specific research programs (e.g., CIAT, CIMMYT, CIP, IRRI) while several centers (e.g., CIAT, CIP, ICRISAT, IRRI) have undertaken more formal ex ante assessments as part of a center-wide, priority-setting process.

Probably the most important and well-developed body of work along these lines is from CIMMYT. Scientists there have developed a global data base on the resources NARS commit to wheat-breeding research and the adoption of modern wheat varieties (including those developed at CIMMYT or based on CIMMYT-bred parental lines), and have undertaken a number of complementary adoption-cum-research-evaluation studies of particular aspects of CIMMYT work in particular countries.

It is perhaps no surprise to learn that CIMMYT (and IRRI) have a reasonably well-developed track record of ex post studies. They are the centers that have been in existence the longest and it takes many years from inception (upwards of 20 to 30 years) before the technology development and adoption process has run enough of its course to make meaningful ex post studies possible. A significant amount of evaluation and priority-setting work has been initiated by other centers over recent years. For example, CIP has recently completed a series of nine (largely externally executed but internally managed) studies of the impacts of particular aspects of its research; ICRISAT has instigated a program of "Research Evaluation Impact Assessment" work within its Socioeconomic and Policy Division; IFPRI has recently formed an Impact Assessment Committee and secured funding support to begin the difficult task of developing procedures to evaluate the effects of policy research.

A more comprehensive, but by no means complete, accounting of recent and ongoing research evaluation work at the centers is given in section 5 and appendix A. Much of this effort is directed and carried out by the socioeconomic units within the centers. The continued commitment of significant socioeconomic resources around the CGIAR to this type of work is remarkable when one contemplates the substantial reduction and turnover in socioeconomic personnel throughout the system in recent years.

External Agencies

A significant body of ex ante analysis of the global consequences of agricultural R&D (including, CGIAR research) has been carried out in the Research Evaluation unit at the Australian Centre for International Agricultural Research (ACIAR). While one can differ with some of the particular assumptions and approaches, the ACIAR work has the advantage of having used a consistent, coherent, and transparent methodology to look at a range of types of research from a multinational perspective. It is the transparency of the method that makes informed criticism possible (and this same transparency permits criticism to be of a form that

leads to refinement in the methods and the value of the results for resource allocation).[18] The problem with the ACIAR work for CGIAR purposes is that it has not been designed to serve the CGIAR objectives (and therefore does not satisfy all CGIAR requirements).

Much of the past research evaluation effort around the system has actually been carried out by external consultants or by staff temporarily seconded to the CGIAR from external agencies. A good deal of independent analysis has also been done by individuals working in universities and elsewhere, much of it funded by agencies such as ACIAR and USAID that are also principal donors to the CGIAR. While this mode of operation has the potential to inject fresh ideas and bring with it a degree of independence, it also has some disadvantages. It is difficult to achieve the necessary standardization of analytical methods, to ensure the studies can be meaningfully compared across institutions at a point in time, or at a particular location over time, and provide an overall picture that is greater than the sum of the parts. And, with a heavy reliance on external analysts, it is difficult to accumulate the knowledge and data, across and within the centers, so that the system benefits by lowering the longer-run costs of doing this type of work. It is not surprising that CIMMYT, the center that has committed the most time of in-house scientists to research evaluation, has the most well-developed (and, well-regarded) body of work to show for its efforts.

Lessons from the Past

Commodity-Based Evaluation: A common misperception is that, because most applications of economic approaches to evaluating R&D are based on commodity-market models, the methods themselves are *intrinsically* limited in their capacity to evaluate important effects of R&D. Particular concerns are expressed about the applicability of those models to evaluate the consequences of R&D for input markets (e.g., labor use) and the environment. Several comments are pertinent (see Alston, Norton and Pardey 1995 for more such arguments):

- with reasonable and meaningful objectives for R&D, the appropriate measures of research effects are economic surplus measures that can be represented in the context of supply and demand models for commodities, inputs, or environmental goods (and economists have demonstrated that the benefits measured in the commodity-market context capture *all* of the economic effects on consumers, resources, and so on, if they are defined with care)
- traditional productivity growth is (and, we argue, should be) the primary objective of most of the CGIAR's R&D, and it is appropriately evaluated using a commodity market model
- resource degradation effects of changing technology *can* (and probably should) be incorporated into conventional productivity measures and thereby can be represented appropriately in a conventional research benefit model

[18] See Davis, Oram and Ryan (1987) and Davis and Lubulwa (1995).

- economists have no problems in incorporating environmental externalities associated with agricultural production in their models in principle; the practical implementation of those models is hampered by inadequate *scientific* information on the technical relationships between agriculture and the environment. This same point holds in relation to the effects of research and technical change in modifying the environmental impacts: if the relationships can be quantified physically, the evaluation is relatively straightforward.

In short, every one of these typical objections is based on a misunderstanding about the applicability of the supply and demand and economic surplus methods for evaluating a broader range of research evaluation issues. The methods, perhaps with minor modifications, are perfectly applicable to many issues that they have not previously been used to address. Finally, the challenge to those who would reject conventional economics is to ask: what is the alternative?

Standardization: The growing use of evaluation methods has developed in an uncoordinated fashion throughout the CGIAR system, and as a result there is major variation in the objectives, data sources and treatments, and analytical methods used (although mostly economic surplus methods are used, and increasingly so). Some such variation is productive, as a way of developing methods, and appropriate as different contexts call for different approaches (reflecting different questions or different constraints of data availability and so on). On the other hand, unnecessary variation is undesirable since it reduces comparability, leads to wasteful duplication of effort in developing and refining methods, and reduces the gains from exploiting economies of size, scale, and scope.

Matching Benefits and Costs: A variety of rates of return might be computed, depending upon the purpose of the analysis. For instance, one might be interested in the rate of return to an investment by a particular group of producers (or by people in a particular geopolitical region -- e.g., a state or a nation), to work conducted by a particular research institute, to a particular program of research, or to a particular component of a research program (e.g., the extension phase). There are important differences between marginal and average rates of return, real and nominal rates of return, and private and social rates of return.

Sometimes the issue of which rate of return is of interest is not defined clearly, leading to a mismatch between the basis upon which the benefits are measured or the basis for measuring the costs, and the basis of interest. Specifically, the rate of return can be computed only by comparing the full costs associated with a defined R&D investment, borne by the group of interest, with the full benefits from that investment going to that same group. Problems arise from mismeasuring the relevant benefits or costs (e.g., by including benefits going to some other group or attributable to some other R&D investment).

A decision must be made about which rate of return to attempt to compute in order to determine which benefits and costs are relevant to include in the analysis, and to determine which counterfactual simulation experiment to conduct in order measure the benefits and

corresponding costs including deciding how to handle the spillover consequences of R&D conducted elsewhere and other "ceteris paribus" conditions.

Adding Up: Some seek to evaluate all programs, perhaps on an ex post basis for showing past effectiveness. Another (but unrealistic) alternative is to evaluate the set of programs in aggregate. It is uneconomic to evaluate everything in sight. An alternative is to evaluate the most important (or most critical) alternatives and acknowledge that some things have been left out. This has the consequence that adding up the benefits from all the things that have been evaluated will understate the total, but it does serve to inform the key tradeoffs. This is an example of different evaluations serving different purposes. And, a representative sample of programs is an effective indicator of the entire set of activity.

Beyond Summary Statistics: The reduction of the evaluation results to summary statistics (such as the rate of return or a benefit-cost ratio) has a virtue of enabling comparisons to be made, but it also involves a loss of other information. Of course, it is not necessary to throw away the other information (e.g., on adoption, yields, price effects, incomes, streams of benefits and costs, environmental externalities and research spillovers) used to compute the summary statistics. Much of this information is intrinsically interesting and provides a more complete description of the research process and its effects, allowing a more complete understanding of the past that assists in making informed judgements about future research resource allocation decisions.

"Alternative" Indicators of Impact: The emphasis in this paper is on placing an economic value on the overall benefits of CGIAR research. Expressing the consequences of R&D in terms of a standardized money metric is a particularly useful way of summarizing large amounts of disparate scientific and economic information that would otherwise be extremely difficult to use in any meaningful way for making more informed funding and resource allocation decisions. In particular, such summary statistics make it possible to systematically explore the tradeoffs involved in decisions about whether to fund research or to provide some other types of development assistance (e.g., fund education or health services or invest in other forms of rural infrastructure) or the implications of targeting R&D to particular geographic regions, particular groups of people, or particular lines of research.

But as a practical matter, financing and resource allocation decisions can be, often are, and will continue to be made on the basis of "impact" indicators other than those that place an economic value on the *overall* consequences of agricultural research. This is so for a variety of reasons. One reason may simply be that the requisite research evaluation studies have not been done -- perhaps due to data, resources, or some other constraints. Shortcut or proxy indicators of impact (e.g., congruency measures) have been used as a substitute for the more complete measures in some cases. While using shortcut measures may be expedient, as Anderson and Gryseels (1991, p. 326) observed, these measures often represent "much less than half the story and it is the conditioning multipliers of real research opportunity that must guide any informed research policy analysis."

Another reason is that some decision makers may have preferences for assessing the benefits (or costs) flowing from research to a particular geographic region or group of people (e.g., South-East Asia versus Africa, or small versus large farms, or even producers versus consumers). In such cases the evaluation effort may need to go beyond accounting for total net benefits and provide information on the likely pattern of benefits as well. In other cases decision makers may prefer to have information on the employment consequences of R&D, the consequences of R&D in terms of the nutritional status of some target group, more direct indicators of the effects of R&D on the environment, and so on. Providing these types of indicators draws from the same set of data used to generate the more aggregate indicators and may simply require reworking the analysis to present the same information in a different way. But in virtually all cases it requires that additional data be collected and analysed.

Some decision makers are satisfied with less formal indications of the consequences of R&D. Information on the overall rates of adoption and more specific details on the rates of adoption by particular groups of farmers, data on the experimental or on-farm yield effects of R&D, or even indicators of R&D "output" such as the number of scientific publications or the number and nature of new varietal releases can be useful ways of demonstrating the potential effects of research. As discussed in the previous section, many of these indicators are precisely the same variables used in carrying out a more formal assessment of the benefits from R&D. If reporting this type of information is deemed useful then it would make sense to do so. But in practice, evaluation efforts that do not carefully collect and use this type of information in a more structured way (drawing on basic economic principles and evaluation methods) can quickly degenerate into an ad hoc compilation of indicators that in truth give little meaningful indication of the links between investing in R&D and the consequences that concern decsion makers and audiences more generally.

4. Principles for Designing Evaluation Institutions

Functions to be Defined

The following related functions are necessary components of research evaluation in the CGIAR system:

- developing and maintaining data bases that will support the various evaluation activities
- developing and maintaining analytical capacity
- developing and defining the standardized methods and procedures for evaluation, for using the results from evaluation for setting priorities, and for meeting accountability requirements
- evaluating research programs and communicating the results to decision makers within the CGIAR and other relevant audiences.

Cost Effectiveness

Any new institutional arrangements should be designed with a view to taking advantage of the existing capacity to perform those functions that exists in the system, so that wasteful duplication is avoided and so that effective current activities are not crowded out. At the same time, it should be recognized that different types of evaluation are likely to continue (optimally) to be carried out at different levels of the system in order to serve different types of decisions and users of the information. And, over time, the appropriate arrangements will evolve: we mustn't be bound by what exists today, or is good for today. Cost effectiveness over time requires built-in flexibility and adaptiveness, defined by principles rather than arbitrary rules that may become redundant quickly.

One important distinction is between long-run and short-run cost effectiveness. Taking a short-sighted view might lead to a decision to conduct ad hoc studies on a tactical basis. But the strategic choice is more likely to involve a significant up-front investment in developing methods, data, and analytical capacity. Once the initial costs have been incurred, the maintenance costs to keep the capacity going (a much more effective capacity than would be achieved by the ad hoc approach), should be less than would be required for the current types of evaluation activity. The individual centers seem to be developing their capacity in a more strategic fashion than the central bodies (TAC, CG Secretariat) that have tended to invest in large-scale, ad hoc, episodic studies rather than operate on a continuing basis, perhaps more modestly, as a long-run perspective would suggest.

Cost effectiveness implies that the information that is produced is demanded by someone for some purpose. Relevance requires timeliness, but it is all too easy to sacrifice substance (content and quality of information) for speed. It would be short sighted to emphasize timeliness so much that the information produced would be uninformative. On the other hand, unless the information is available in time to be used, it is useless. Optimizing this tradeoff is an important element of the institutional design. A flexible structure may be able to provide rougher estimates on a more timely basis while providing better answers, later.

Independence versus Consistency

What is best for a particular center may not be best for the system as a whole. Two important aspects to consider are incentive compatibility and independence. Individual scientists, or centers, or even the CGIAR system as a whole, may be unable to provided a disinterested perspective. With care, evaluation institutions can be designed to minimize the degree of subjectivity and possible bias in results. One advantage of greater transparency, standardized methods, and more formal quantitative approaches is that potential biases are more likely to be exposed.

System-wide impact studies may be desirable for accountability reasons, in order to shore up funding support from outside the system. Studies like the TAC scoring study, that involve no objective evaluation, offer little value in this regard. Any internal study may be open to

skepticism since the CGIAR has clear incentives to overstate benefits, but this skepticism can be mitigated if transparent and objective methods are used, preferably on a consistent basis over time. The use of external consultants may yield the appearance of greater objectivity (but often the appearance may be overestimated, since successful consultants usually establish early on what answer the client is looking for, with an eye on the odds of repeat business). On the other hand, some consistency and continuity may be lost if consultants are used rather than a permanent capacity within the CGIAR.

A particular virtue of keeping the evaluation activity within the system is that, the continuing involvement of the scientists in the evaluation and priority-setting activities fosters the development among them of an economic way of thinking about their work. When consultants are engaged to substitute for in-house reviews, this engagement of the scientists in the activity should not be forsaken. It would be well if donors, also, could become involved within the process of evaluation, to understand what it means, rather than simply wait and hope to digest the results.

5. Current Capacity

Before committing significant resources to revamping and, perhaps, expanding the CGIAR's research evaluation effort, it is worth taking stock of the existing research evaluation capacity. Here the focus will be on the analytical and decision-support capacity within the CGIAR. Expertise and data sources outside the system is currently being, and should continue to be, tapped. Time and resource constraints on the present study preclude us from comprehensively assessing the capacity outside the CGIAR. The same is partly true of our assessment of the resources within the CGIAR. However, although we do not claim to have a complete stocktaking of the internal CGIAR capacity for research evaluation, it is likely that we have identified the major elements of it.

Human Capital

One of the real, binding constraints for doing quality research evaluations is the time and expertise of the scientists committed to the work. A quick, subjective assessment suggests that, compared with the past, the CGIAR as a whole is currently reasonably well-endowed with analytical capacity for conducting research evaluation studies. This is not to say that current investments in this area are adequate or that all centers or system-wide decision-making bodies (e.g., TAC, CG Secretariat) have sufficient access to this type of expertise. But it is true that despite (or, maybe, in response to) recent cutbacks in financial support, an increasing number of centers have committed significant staff time to this line of work, and others have made plans or are actively gearing up to do so. To the authors' knowledge, the current state of play is as follows:

TAC Secretariat: The quantitative research evaluation/priority-setting work undertaken as a part of the last round of strategic priority-setting deliberations, led by TAC, is discussed in section 2 of this report. Most of the personnel TAC assembled for that exercise have moved on and current secretariat personnel have very little time (if any) to provide this type of decision

support to TAC. It is unclear to what extent the data sets assembled and methods developed in the TAC study are actively being used by centers within the system.

CIAT: In September 1991, CIAT established an Impact Assessment Unit within the Office of Director General. The activities of this Unit built directly on the prior work at the institute to quantitatively evaluate the size and distribution of the benefits from improvements in bean technologies. The Unit led recent efforts within CIAT to undertake a comprehensive, ex ante assessment of the Center's programs and projects as part of a priority setting exercise in support of the development of CIAT's 1994-98 medium-term plan. This ex ante assessment effort is on-going and the adoption cum impact assessment work has moved beyond bean technologies to include cassava and ley-farming systems.

CIP: The economics group at CIP, in conjunction with colleagues in collaborating NARSs and universities, recently completed drafting a report titled *Case Studies of the Economic Impact of CIP-Related Technologies* that contains nine studies of the impacts of CIP research related to genetic improvement, IPM, and seed propagation technologies.

CIMMYT: The Economics program at CIMMYT continues to commit significant staff time and resources to developing new data sets, as well as analyzing and reporting on the impact of the center's work in some depth in selected countries and more broadly on a regional level. Most of the effort is directed to ex post assessments. Although a good number of these studies were done in collaboration with colleagues in other agencies, a feature of past CIMMYT work in this area was the commitment of significant CIMMYT staff time to executing (not simply managing) such studies. Unfortunately the resources committed to this line of work have been reduced as a result of staff turnover and in line with more general cutbacks throughout the center.

ICARDA: ICARDA undertakes four different types of impact assessment: (a) technology adoption and impact studies; (b) institutional impact and development studies; (c) regional and subregional sector studies; and (d) technology evaluation. The first two of these focus on the impact of past work by the Center (ex post assessment), whereas the second two address the issue of potential impact (ex ante assessment). The technology focus of impact assessment is on sustainability and productivity of ICARDA's mandated crops and farming systems at the regional, national, and farm levels, and includes an assessment of the distributional consequences of research. The institutional concerns are with the effectiveness of the Center's relationship with NARSs of the WANA region and with the progress of technology transfer to ICARDA clients and beneficiaries.

A number of individual adoption and impact studies and technology evaluation studies are conducted each year by ICARDA staff in partnership with NARSs scientists. These are an integral part of ICARDA's work in technology generation and transfer. Institutional impact and sector studies take place less frequently and are usually undertaken in response to specific special project and/or donor needs.

ICLARM: As a small center with a wide mandate and limited resources, priority setting is a critical factor in how ICLARM decides its work program. With the additional limitation that the majority of ICLARM's resources are provided as project monies, it is imperative that research and related work programs define the work agenda and that it is not driven solely by the availability of earmarked project monies.

ICLARM's past priority setting work was centered around the formulation of ICLARM's 1994-98 Medium Term Plan, which set the agenda for the subsequent choice of projects for additional (compementary) funding. In April 1995, the ICLARM Board of Trustees approved a program of impact assessments of ICLARM work based, in the first instance, on a wide range of case studies from which general principles regarding the possible impact of future ICLARM activities are to be drawn. The assessments will be commissioned by ICLARM but, where possible, undertaken by independent researchers. Their purpose is threefold: (a) to demonstrate the value or otherwise of ICLARM's previous work, (b) to guide the choice or further research directions into areas of best return, and (c) to value research on aquatic resources generally and thereby strengthen the case for greater support at the national as well as international level.

Prior to this plan of assessments, ICLARM undertook extensive scientometric research on its own publications (publication and citation patterns and rates) and those of NARSs and other CGIAR centers. ICLARM also maintains records of the scientific use of its computer software.

ICRAF: The Characterization and Impact Program is the newest of the four research programs at ICRAF. It was created at the beginning of 1993, following the restructuring which occurred when ICRAF joined the CGIAR system in 1991. Before then, and during the period 1986-1990, work of relevance to the Program was undertaken in both the Collaborative Programs Division and the Research Development Division. Emphasis was placed on diagnosing land-use systems and predicting agroforestry potentials in these systems, largely on the basis of hypotheses and non-quantitative data.

In July 1990, these two Divisions merged into one Research Division that included two programs relevant to the objectives of today's Characterization and Impact Program: (a) the Environmental Characterization and Analysis and (b) the Policy, Adoption and Impact Programs. The Environmental Characterization Program had only one part-time scientist and did not receive a high priority. The Policy, Adoption and Impact Program focused on the on-farm testing of agroforestry technologies and the use of participatory trials. In addition, some diagnostic work was undertaken on shortages of wood products, farmers' soil fertility management practices, and farm nutrient budgets, and in 1992 land tenure studies were initiated in four African countries. At the beginning of 1993, the two programs were merged into the current Characterization and Impact Program.

ICRISAT: A Research Evaluation and Impact Assessment initiative led by the Socio-Economics Division has been underway for several years. ICRISAT invested considerable effort in applying a scoring model approach to assist in developing its 1994-98 medium-term plan.

This work has now broadened to include a well conceived and integrated set of research evaluation cum priority-setting activities. This involves representative examples of ex post evaluations of ICRISAT research, adoption studies, and the institutionalization of a quantitative ex ante research evaluation capacity within ICRISAT to support its research decisions.

IFPRI: In October 1994 work under the auspices of the IFPRI multi-country program on Agricultural Research, Extension, and Education Policy (MP4) was initiated. The program has developed quite rapidly over the past several months. In addition (and complementary) to its broader program of policy research, MP4 will be directing a considerable amount of its attention to evaluation and priority setting issues. The intent is to continue developing and extending the basic research evaluation methodologies to deal with a broader range of problems (e.g., cross-commodity effects, research spillovers, and research aimed at ameliorating the resource degradation and adverse environmental effects of agricultural production.) Program staff are supporting an IFPRI-wide effort aimed at assessing the effects of policy research and are also working to stimulate an integrated, systemwide program of research on agricultural genetic resources policy, which has a significant evaluation and priority-setting component.

IFPRI will also work with external counterparts in a number of selected country studies. The primary objectives of these studies are to field test new evaluation methods, generate results that have intrinsic value in the policy formulation process, and accelerate the transfer of these policy analysis technologies to national systems. A significant effort to develop and operationalize research evaluation cum priority-setting methods that support policies and strategies for regionalized programs of research will be undertaken, with special emphasis, at least initially, on Latin America and the Caribbean. Finally, IFPRI will continue to collaborate closely with ISNAR, other CGIAR centers, and NARS partners to maintain a global data base of investments in agricultural research. They will also undertake broad (i.e., cross-country) policy assessments of these data in ways that will be directly useful for formulating research investment strategies. A distinguishing feature of IFPRI's work is the on-going effort to place these research investment decisions in a broader institutional and policy framework and assess the options and tradeoffs involved in committing public resources to R&D.

ILRI: ILRI has completed several studies that involve impact assessment methods. Work carried out under the auspices of the program on forage legumes in the sub-humid zone of West Africa identified fodder availability as a major constraint to increased livestock production. Follow up work suggested that "fodder bank" technology was an appropriate response to this constraint and efforts were made to monitor the adoption of this technology throughout West Africa. Additional studies have evaluated the benefits from adopting "broad bed maker" technology as well as the impact of livestock diseases and the likely economic benefits to controlling them.

Current and planned future work in this area involves an integrated set of studies to evaluate the environmental consequences of alternative trypanosomiasis control measures as well as an assessment of the benefits and costs of improving livestock production under trypanosomiasis risk in West Africa. ILCA staff are also reviewing the available impact

assessment methodologies with a view to adapting and applying these methods to the problems of evaluating livestock-related technologies.

IPGRI: In February 1993 IPGRI initiated an impact assessment project. The project is addressing a number of issues including an assessment of the impact of IPGRI's activities on genetic resources conservation and use, several case studies on the use of land races and wild species in breeding modern varieties, and the development of criteria and indicators that can assist IPGRI in evaluating its own projects and programs. Most of this work involves ex post evaluation approaches but a significant effort is being made to identify ex ante indicators that will assist IPGRI in formulating its own priorities.

IRRI: The Economics program at IRRI continues to do a significant amount of work on adoption studies, the environmental consequences of rice production technologies, and some ex post assessments. In recent years, increasing emphasis has been given to understanding the extent and value of varietal transfers (i.e., research spillovers) and to applying more formal methods to developing IRRI's priorities for rice research (see Evenson 1994, for example). An effort to systematically compile data on rice research in NARSs was instigated within the past year.

ISNAR: ISNAR has engaged in two main types of impact assessment: (a) assessment of the impact of agricultural research on economic efficiency, equity, and sustainability, and (b) assessment of its own impact on the institutional development of NARS. Ex ante and ex post assessments of agricultural research impact have applied scoring and economic surplus models in a number of countries. Considerable attention has been paid to the integration of economic analysis and geographical information systems, in order to estimate impacts in different agro-ecological zones and the spillover of research benefits among zones. This work has culminated in the publication of a book on the state of the art in this field (Alston, Norton, and Pardey, 1995).

In 1991, ISNAR prepared an in-house assessment of its own achievements, impacts, and constraints, in preparation for its last External Review. This assessment was based on a review of ISNAR's research, training, and advisory outputs, a series of country-level case studies, and staff assessment of results of collaboration in each country and region. More recently, six NARO leaders have prepared brief personal assessments of the impact of ISNAR's work with their organizations. A series of more formal case studies of ISNAR's impact are now being designed, and will be carried out over the next three years. The main purpose of these studies is to draw lessons from past experience, in order to improve ISNAR's future institutional development efforts.

ISNAR prepares training materials and organizes training events on economic impact assessment and on evaluation more broadly. An approach is also being developed to assess the performance of agricultural research organizations, based mainly on scoring indicators of research outputs. These efforts are aimed to strengthen the capacity for planning, evaluation, and impact assessment in NARSs. For example, since 1992, ISNAR has worked with a group

of professionals and organizations in Latin America and the Caribbean, to prepare training materials and form a group of trainers specialized in planning, monitoring, and evaluation of agricultural research. This work will continue until 1997 in LAC, and similar efforts are being planned for other regions.

Data

To do meaningful research evaluation studies (whether they be ex post or ex ante) requires a basic set of market data on the prices and quantities of production, an assessment of the yield increasing or cost reducing effects of past, present, or possible future research, an understanding of the markets being modeled, and a clear identification of the cost involved in carrying out the research and getting it adopted. While the data requirements to undertake this work are well known to a competent analyst, the level of detail deemed appropriate, the identification of data sources, and the means and methods used to process these data are important considerations that are in some senses particular to each study. Nevertheless, there are some primary data sets that are common to most studies, and the CGIAR could reap significant economies of scale and scope from collecting and reporting these data in a coordinated rather than ad hoc fashion.

Germplasm Data: Tracking the parentage, development, release, adoption, and local productivity performance of new (CGIAR) germplasm is necessary for evaluating the economic consequences of the system's significant investment in genetic improvement research. Developing standardized procedures for compiling and reporting these data, as part of the SINGER initiative, will substantially reduce the costs and improve the value of using this type of data for research evaluation and research policy purposes.

Market-Related Data: A substantial number of international and national agencies compile and report data on population, price, consumption quantities, production quantities, and so on that is required for research evaluation. Difficulties arise in gaining ready access to updated versions of these data and compiling them in ways that are useful for research evaluation purposes (particularly, obtaining spatially disaggregated and geo-referenced versions of these data so they may be incorporated into geographical information systems where it is necessary to do so). For more detailed evaluation studies it may also be necessary to measure the nature and size of market distortions. Currently some centers (e.g., IRRI, CIMMYT, ICRISAT, and CIAT) routinely compile and report such data and they are becoming increasingly available via the INTERNET. And a variety of other institutions are compiling consistent data on market distortions (e.g., OECD, USDA). IFPRI is probably the primary repository of this type of information within the CGIAR although a concerted effort would be required to process this information in a way that is useful for research evaluation purposes.

Agroecological Data: Because the impact of much agricultural research is heavily influenced by agroecological factors (e.g., soil type and slope, elevation, rainfall, and length of growing period) it is important for evaluation purposes to incorporate such information into the analysis. This was the intent underlying TAC's efforts to introduce an ecoregional concept in

to its last priority setting exercise. This type of information is extremely valuable in assessing the site specific impacts of agricultural R&D, including an assessment of the potential for research to tackle some of the natural resource degradation consequences of agriculture. CIAT has probably the most well developed GIS capabilities within the CGIAR although other centers are strengthening their capabilities in this regard. And as the UNEP/GRID initiative develops further, GIS information will become increasingly accessible throughout the system. But, as with the market-related data, some significant effort will be required to process this information in ways that are useful for research evaluation purposes. .

R&D Costs: Meaningful research evaluations require matching research benefits to research costs. For CGIAR research, which is often carried out in conjunction with, or is further developed by the efforts of, national research systems, it is necessary to have an appropriate classification of CGIAR costs as well as the research cost incurred by NARS partners. The CG Secretariat and individual centers compile data related to the costs of R&D in the system although some significant work would be required to develop and implement standardized reporting procedures throughout the system to make these data really useful for research-evaluation purposes.

ISNAR has sustained a long-running effort to develop a comprehensive set of internationally comparable data on investments in agricultural research by national systems. ISNAR and IFPRI have recently proposed a system-wide initiative to further develop and institutionalize this undertaking within the CGIAR in ways that would directly support research evaluation efforts at individual centers and for the system more broadly.

Integrated Voice and Data Network: There are a number of data bases that are common to most research evaluation exercises. This includes information on commodity prices, quantities produced and consumed, price and income elasticities, research costs of national systems, and so on. While a good deal of these data are available from international agencies such as FAO and the World Bank (and many of them are maintained at IFPRI) a significant amount of work must be done to standardize these data sets and put them in a form that makes them readily usable for research evaluation purposes. And, increasingly, disaggregated data are available from national statistical agencies and access to these data would significantly improve the quality and coverage of research evaluation efforts.

Although there could be significant scale and scope economies from coordinating efforts throughout the CGIAR to collect, process, and store this information, at present there are no resources earmarked for such work nor the institutional mechanisms in place to do it. But the integrated voice and data network currently being installed throughout the CGIAR will make it feasible to launch a more concerted effort to make "standardized" data sets available to each center in the system and for that matter NARS collaborators who have access to the INTERNET.

Methods

A variety of formal and informal methods have been used for both ex ante and ex post assessment of agricultural R&D projects and programs. Some of the methods are totally ad hoc, and without foundation in economics (or any other respectable field related to evaluation as a science). This comment relates, perhaps, especially to the more informal evaluations, many of which are not even recorded. For the most part, however, the recorded results of formal R&D evaluation (and priority-setting) activities relate closely to conventional cost-benefit analysis, using conventional (and widely accepted) economic measures of economic impact. (Of course non-economic measures are required for non-economic impacts, such as physical consequences of new varieties for yields and adoption and soil erosion, say. But often these so-called non-economic impacts are really most interesting when translated into economic measures -- such as the benefit from higher yields in terms of farm incomes, lower consumer prices, the value of soil losses, and so on.)

The formal methods of cost-benefit analysis are well documented in various texts. Some special issues arise in applying those methods to evaluation of agricultural R&D projects (due to differences between general projects and typical agricultural R&D projects). Agricultural economists have, over the past thirty years or so, developed approaches for dealing with many of the special issues, most of which concern translating the effects of research-induced changes in productivity into shifts in supply (or demand) and, from there, into measures of economic benefits to particular interest groups. Models have been developed to deal with international trade in commodities, research spillover effects, multimarket impacts in multistage production models, effects of R&D in the context of markets distorted by government programs, environmental externalities, risk, and imperfect competition. Many of the issues have been resolved in principle, but some in-principle issues remain unsolved as yet. Some of the issues that have been solved in principle are still difficult to deal with in practice.

One of the foundational studies in this area was by Edwards and Freebairn (1981), elaborated further in the work by Davis, Oram and Ryan (1987). These studies dealt explicitly with international trade. More recently, Alston (1991) reviewed and extended that literature to deal with multimarket R&D effects more generally. This month, the book *Science under Scarcity* by Alston, Norton and Pardey (1995) was published for ISNAR by Cornell University Press. This book documents in considerable detail the various methods that have been developed for agricultural R&D evaluation and priority setting. In addition to the book itself, the authors have developed a software program (*DREAM*) that implements standardized approaches from the book, in a user-friendly, menu-driven computer package. That package will be available, along with other hands-on tools, in a set of training materials being developed as a companion to the book and the computer program.

In addition to this set of materials, a number of other independently developed lines of work have also developed manuals, rules of thumb, and computer packages. Many of these packages and rules of thumb are incomplete and potentially misleading. Most are not as comprehensive in approach as the *Science under Scarcity* cum *DREAM* package. Work is

currently under way to evaluate these alternative packages for R&D evaluation and priority setting.

Finally, as partly addressed by Alston, Norton and Pardey in *Science under Scarcity*, some alternatives to formal cost-benefit evaluation have been proposed and used by some. We would argue that those alternatives are potentially misleading; hence, when shortcuts are necessary, it is best to attempt to approximate the results of cost-benefit analysis.

6. Policy Options

In writing this report, we have taken as given that there will be some change in the CGIAR's approach to research evaluation (and the use of the results from evaluation for priority setting). A new institution will be developed. The important questions that remain are what roles should the new institution play, including how large a role (perhaps defined in terms of the total funding support).

Among the roles should be a policy advisory role. One purpose of developing a new institution is to provide a capacity for developing more complete answers to a number of issues raised in this report, answers that may change over time as the system evolves and matures. Some of that evolution, it is hoped, would be influenced by the information developed by a systematic process of research evaluation and system assessment. Whether the potential improvements in economic efficiency are realized will also depend on the details of the institutional arrangements affecting the management of research resources (in terms of incentives and procedures for allocating resources). That is, like other research, research evaluation research does not end with the evaluation itself; it is not useful until the results are used in decisions.

Resources for Evaluation

How much should be spent on evaluation? This is a difficult question since the benefits from research on evaluation of research are relatively hard to estimate. Suppose the CGIAR system spends $300 million a year and generates a benefit-cost ratio of 2:1, equivalent to a benefit of $600 million this year from this year's research. Then, if an investment in evaluation would lead to a one percent increase in total program support ($3 million), there would be a gross global research benefit of $6 million and a *net* global research benefit of $3 million against which to charge the costs of the evaluation. In other words, from the point of view of global economic efficiency, it would be worth spending up to $3 million per year on research evaluation if the results of such evaluation would generate an additional one percent in funding. Alternatively, keeping the total funding constant, it would be worth spending $3 million per year on research evaluation if it would result in a half of one percent improvement in allocative efficiency (i.e., in total research benefits). Even very modest expectations about the potential for research evaluation work to lead to either increased funding, or increased allocative efficiency, would be needed to justify an annual budget of $1 million or more.

Another way to look at the same issue is to consider the question: what could we expect if we did no evaluation? It might not be realistic to expect a continuation of the existing funding without any concessions to donor demands for accountability. The recent experience shows that reductions in total funding of $10 million per year are possible. If an effective research evaluation institution could stave of a budget cut of $10 million per year, it would be worth investing $10 million per year in evaluation (assuming a benefit cost ratio of 2:1 for the research).[19]

Main Alternatives

Various approaches may be used to promote the evaluation role within the CGIAR system. These alternatives range from providing incentives for individual scientists or programs to conduct evaluations on their own behalf through to center-based activities. It seems likely that the new institutional arrangements adopted by the CGIAR will involve several elements as described below. The crucial aspect of the design may be how to combine the elements to exploit their individual strengths, and to capitalize on the complementarities and synergisms that can be achieved in a coordinated approach.

Competitive Grants: A number of institutions allocate competitive grants among scientists based on their assessment of the likely outcome from the research, including an ex ante economic evaluation. The CGIAR (or individual donors) could promote evaluation activities by making funds available for allocation among scientists, centers, or programs based upon this type of competition. It would seem to be worth trying this on a small-scale basis, at least initially.

Tenders: Similarly, but as a separate notion, the CGIAR could invite individual scientists, centers, programs, or external agencies to tender for any or all of the evaluation functions it envisages.

Ad Hoc Consultants: As discussed above, consultants have certain advantages and disadvantages relative to in-house capacity. Major use has been made of external consultants for past ad hoc, system-wide and center-based reviews.

Strategic Alliances: An alternative to the development of solely in-house capacity, or the use of ad hoc consultants, is the development of a long-term relationship with individuals or institutions external to the CGIAR. Such strategic alliances could have the best of both worlds. They would complement the in-house effort while providing both continuity cum consistency (unlike ad hoc consultants) and a greater degree of independence (unlike internal arrangements).

Center-Based Activities: Centers are the natural place to continue to carry out ongoing

[19] Between 1992 and 1993 the total contribution of donors to the CGIAR declined by about $7.5 million -- from $318.7 million to $311.27 million (Unpublished CGIAR budget documents 1994). The amount actually contributed in 1993 was substanially below the amount expected, based on an extrapolation of past funding trends.

evaluation activities related to the intra-center allocation of resources. They could also provide much of the evaluation of center-wide performance needed for cross-center comparisons at the CGIAR system-wide level, or for representation of the system and the centers to donors. The issue to resolve may be where to divide the roles between center-based evaluation and centralized work, given the differing information requirements for different decisions. Such choices should allow for the fact that centralizing the activity is only one way of standardizing procedures and ensuring data bases are maintained.

To the extent that any change in the institutions entails increased expectations about the research evaluation outputs that would be produced by center-based social scientists, there must be either a corresponding decision to give up some current outputs from them or to provide additional resources. The latter choice would seem to be more sensible, given both (a) the relative smallness of the socioeconomic resource base in the centers and the implications for their programs of asking those people to switch to evaluation and (b) the fact that not all socioeconomists have the requisite expertise. This is especially so in relation to center-based support for system-wide evaluation.

Centralized Capacity: Centralized capacity has advantages, in relation to economies of size, scale, and scope and standardization of measures for comparison to be traded off against disadvantages of remoteness from the actual research activity being evaluated. Some decisions necessarily cannot be served well simply by relying on center-based evaluation (e.g., system-wide evaluations). Also there is a role for some centralized collection of data and development of procedures and standards. The difficult distinction, as noted above, is between the different roles of center-based versus centralized evaluations. Perhaps the difficulty of choosing could be taken to mean that for many such choices, the choice does not matter very much and expediency, the opportunity to build upon and exploit current capacity in the centers, may be a relevant criterion. These issues cannot be addressed easily in principle. The answers ought to depend on facts such as the details of the questions to be answered by analysis, and the characteristics and preferences of the people and arrangements already in place.

In addition to any permanent centralized role, in the near term there is likely to be a requirement for a centralized role in managing the transition, coordinating the changes that are to take place at the centers, and to facilitate communication among the centers which is where we perceive the main part of the CGIAR evaluation activity will reside in the longer term.

Leadership Roles in Implementing Change

Most of the specific questions about optimal design of research evaluation institutions cannot be addressed in a paper of this type. The important point, that we have stressed throughout, is that the design should pay attention to the strengths of the existing institution, the multiple purposes to be served by evaluation, the gains from economizing through coordination, and the virtue in flexibility to adapt when the questions or the circumstances change.

In the light of these observations, we can conclude that the actual development of the institutions should be led by an individual (or very small committee) with

- the authority and leadership qualities necessary to command the respect and cooperation of managers and scientists within the CGIAR
- the detailed understanding of the decision-making and political environment to know what purposes evaluation can and must serve
- the detailed expertise in the evaluation methodologies necessary to design and implement a system that will be able to play the long-term strategic role that we envision.

This activity could well be a part-time role for the person or people identified, even during the development and implementation phases of the institutional change. Later it might be reduced to a convening responsibility administered by the chairman of the CGIAR (or, perhaps, delegated to the chair of TAC, the secretariat, or even rotating among center directors).

Research evaluation, if it is to be worthwhile, is an expensive and time-consuming process. It would unwise to anticipate instant results. It is desirable to proceed with despatch but, more important, to proceed with care. We have seen the evidence from past experience in the CGIAR, that the cheap fix is likely neither to be cheap, nor to fix anything. In taking the next steps, it is certainly desirable to get some evaluation answers for the donors as soon as possible (and a mixed portfolio, of quick and dirty as well as careful and detailed analysis, will surely be part of the design). However, care is needed to avoid taking a path in pursuit of short-term expediency, that precludes the longer-term strategic gains. At the same time, too much consultation can be counterproductive. A good compromise may be to earmark the resources, define clearly what is desired, delegate the authority to the right people, and get out of the way.

7. References

ACIAR. "Strategy for Resource Allocation to IARCs." Canberra: Australian Centre for International Agricultural Research, mimeo., 1994.

Alston, J.M. "Research Benefits in a Multi-Market Setting: A Review." *Review of Marketing and Agricultural Economics* Vol. 59, No. 1 (April 1991): 23-52.

Alston, J.M. and P.G. Pardey. *Making Science Pay: The Economics of Agricultural R&D Policy*. Washington D.C.: AEI Press (forthcoming 1995).

Alston, J.M., Anderson, J.R. and P.G. Pardey. "Perceived Productivity, Foregone Future Farm Fruitfulness, and Rural Research Resource Rationalization." Invited plenary paper for theme VII, *National and International Research and Technology Transfer*, Twelfth International Conference of Agricultural Economists, Harare, August, 1994.

Alston, J.M., G.W. Norton, and P.G. Pardey. *Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*. Ithaca: Cornell University Press, 1995.

Anderson, J.R., R.W. Herdt and G.M. Scobie. *Science and Food: The CGIAR and its Partners*. Washington D.C.: World Bank, 1988.

Davis, J. and A.S.G. Lubulwa. "Integration of Research Evaluation Analysis into Research Institution Decision-Making: An Overview of Progress at ACIAR." Contributed Paper, Australian Agricultural Economics Society annual conference, Perth, February 1995.

Davis, J., P.A. Oram and J.G. Ryan. *Assessment of Agricultural Research Priorities: An International Perspective*. ACIAR Monograph No. 4. Canberra: Australian Centre for International Agricultural Research, 1987.

Edwards G.W. and J.W. Freebairn. *Measuring a Country's Gains From Research: Theory and Application to Rural Research in Australia*. A Report to the Commonwealth Council for Rural Research and Extension. Canberra: Australian Government Publishing Service, 1981.

Evenson, R.E. *Setting Priorities for Rice Research*. draft manuscript, New Haven: Yale Economic Growth Center, 1994.

Gryseels. G. and J.R. Anderson. "International Agricultural Research." chapter in P.G. Pardey, J. Roseboom, and J.R. Anderson, eds. *Agricultural Research Policy; International Quantitative Perspectives*. Cambridge: Cambridge University Press, 1991.

Gryseels. G., C.T. de Wit, A. McCalla, J. Monyo, A. Kassam, E. Craswell, and M. Collinson. "Setting Agricultural Research Priorities for the CGIAR." *Agricultural Systems* vol. 40 (1992): 59-103.

Horton, D., P. Ballantyne, W. Peterson, B. Uribe, D. Gapasin and K. Sheridan (eds). *Monitoring and Evaluating Agricultural Research.* Wallingford U.K.: CAB International, 1993.

Huffman, W.E., and R.E. Evenson. *Science for Agriculture: A Long-Term Perspective.* Ames: Iowa State University Press, 1993.

McCalla, A.F. "Agriculture and Food Needs to 2025: Why We Should be Concerned." CGIAR Sir John Crawford Memorial Lecture, Washington, D.C., October 1994.

McCalla, A.F. and J.G. Ryan. "Setting Agricultural Research Priorities: Lessons from the CGIAR Study." *American Journal of Agricultural Economics* 74 (December 1992): 1095-1100.

Norton, G.W., P.G. Pardey and J.M. Alston. "Economic Issues in Agricultural Research Priority Setting." *American Journal of Agricultural Economics* 74 (December 1992): 1089-1094.

Özgediz, S. "Strengthening Evaluation in the CGIAR: Needs and Options." Draft paper, CGIAR Secretariat, revised March 1995.

J.G. Ryan (ed.), *Building on Success: Agricultural Research, Technology, and Policy for Development.* Canberra: Australian Centre for International Agricultural Research, 1987.

TAC. *A Possible Expansion of the CGIAR.* Report prepared by the Secretariat of the Technical Advisory Committee for International Centers week. Rome: TAC, September 1990.

Tomasi, T.G-. "Sustainability: Concepts and Implications for Agricultural Research Policy." chapter in P.G. Pardey, J. Roseboom, and J.R. Anderson, eds. *Agricultural Research Policy; International Quantitative Perspectives.* Cambridge: Cambridge University Press, 1991.

Tribe, D.E. *Doing Well by Doing Good.* Leichardt, NSW: Pluto Press Australia Ltd in association with The Crawford Foundation for International Agricultural Research, Parkville, 1991.

Tribe, D.E. *Feeding and Greening the World: The Role of International Agricultural Research.* Wallingford: CAB International, 1994.

Wood, S. and P.G. Pardey. "Methods for and Limitations of Ecoregional Analysis." Paper prepared for IFPRI Ecoregional/2020 Vision Workshop, Airlie House Conference Center, Virginia, November 1994 (revised March 1995).

8. Appendix

The references recorded below represent the efforts of the centers over the past several years in relation to research evaluation and priority setting. The listing is by no means comprehensive and the severe time constraints in preparing this document meant that not all centers are represented. That said, however, the listing does show the considerable amount of work done by the centers to document and evaluate the impact of their research and that of their partner NARSs as well as further developing the analytical methods and application tools required to get the job done. A significant number of these publications are in externally peer reviewed journals and books, thereby attesting to the professional quality of this body of work.

CIAT

Gottret, M.V., G. Henry, and M.C. Duque. "Going Beyond Dichotomous Adoption: The Case of Cassava Technology in the Atlantic Coast of Colombia." CIAT, Colombia, draft mimeo, March 1994.

Janssen, W. and N. Ruiz de Londoño. "The Modernization of a Peasant Crop in Colombia: Evidence and Implications." *Agricultural Economics*, Vol. 10 (1994): 13-25.

Cadavid, J.V. and J. Smith. "Adoption of Ley-Farming Systems in the Colombian Llanos." CIAT, Colombia, no date.

Janssen, W., S.M. Teixeira, and M. Thung. "Adoçao de Cultivares Melhoradas de Feijão nos Estados de Goiás, Minas Gerais, Espírito Santo e Rio de Janeiro." *Revista da Economia e Sociologia Rural*, Brasilia, Brasil. Vol. 30 (1993): 321-338.

CIAT. *CIAT's Action Plan. Supplement E: A Preliminary Assessment of the Expected Impacts of Project Outputs*. CIAT, Cali, Colombia, November 1993.

Janssen, W., C.A. Luna, and M.C. Duque. "Small Farmer Behavior Towards Bean Seed: Evidence from Colombia." *Journal of Applied Seed Production*, Vol. 10 (1992): 43-51.

Janssen, W., C. Crissman, G. Henry, M. Lopez Pereira, L. Sanint, and T. Walker. *CIAT, CIMMYT, and CIP: Their role in agricultural research in Latin America and the Caribbean.* CIAT, Cali, Colombia, 1992.

Janssen, W. and S.M. Teixeira. "The Adoption of Improved Bean Varieties in Brazil." in L. Sanint (ed). "Trends in CIAT Commodities 1992." Working document, CIAT, Cali, Colombia, 1992.

Janssen, W., N. Ruíz de Londoño, J.A. Beltrán, and J. Woolley. "OFR in Support of Varietal Diffusion: Beans in the Department of Cajamarca." in R. Tripp (ed). *Planned Change in Farming Systems: Progress in On-Farm Research*, Chichester: John Wiley and Sons Ltd, 1991.

Ruíz de Londoño, N. and W. Janssen. "Un Caso de Adopción de Tecnología: La Variedad de Fríjol Gloriabamba en Perú." *Documento de Trabajo No. 61*, CIAT, Cali, Colombia, 1990.

Sere, C. and L.S. Jarvis. "The Betting Line on Beef: Ex Ante Estimates of Improved Pasture Research for the Latin America Tropics." in *Trends in CIAT Commodities*. Cali: CIAT, 1984.

Pachico, D., J.K. Lynam, and P.G. Jones. "The Distribution of Benefits From Technical Change Among Classes of Consumers and Producers: An Ex Ante Analysis of Beans in Brazil." *Research Policy* Vol. 19 (1987): 279-285.

CIP

Walker, T. and C. Crissman (eds). *Case Studies of the Economic Impact of CIP-related Technologies*, Lima: CIP, draft, January 1995.

Variety Case Studies

Economic Impact of High-yielding Late Blight Resistant Varieties in the East and Central African Highlands
International Potato Center (CIP)

Economic Impact of CIP-24 in China
Song Bofu, Tian Weiming, Wang Jimin Wang Chunlin, Yan Zhengui, and Wang Shengwu

Economic Impact of the High-Yielding, Late Blight Resistant Variety Canchan-INIAA in Peru
Cristina Fonseca, R. Labarta, A. Mendoza, J. Landeo, and T. Walker

IPM Case Studies

Economic Impact of IPM Practices on the Potato Tuber Moth in Tunisia
K. Fuglie, H. Ben Zalah, M. Essamet, A. Ben Temime, and A. Rahmouni

Economic Impact of Managing Sweetpotato Weevil *(Cylas formicarius)* with Sex Pheromones in the Dominican Republic
P. Alvarez, V. Escarraman, E. Gomez, A. Villar, R. Jimenez, O. Ortiz, J. Alcazar, and M. Palacios

Economic Impact of IPM Practices on the Andean Potato Weevil in Peru
O. Ortiz, J. Alcazar, W. Catalan, W. Villano, V. Cerna, H. Fano, and T. Walker

Seed System Case Studies

Impact of Agricultural Research: A Seed Potato Project in Tunisia
Douglas Horton, Roger Corlbaoui, H. Hattab, and Anibal Monares

Economic Impact of the Rapid Multiplication of High-Yielding, Late Blight Resistant Varieties in Dalat, Vietnam
Nguyen Van Uyen, Truong Van Ho, Pham Tung, Peter Vander Zaag, and Thomas Walker

Economic Impact of True Potato Seed on Potato Production in India
V.S. Khatana, M.D. Uphadya, A. Chilver and C.C. Crissman

Fuglie, K.O. "Measuring Welfare Benefits from Improvements in Storage Technology with an Application to Tunisian Potatoes." *American Journal of Agricultural Economics*, Vol. 77, No. 1 (February 1995): 162-173.

CIMMYT

Maredia, M.K. and C.K. Eicher. "The Economics of Wheat Research in Developing Countries: The One Hundred Million Dollar Puzzle." *World Development*, Vol. 23, No. 3 (March 1995): 401-412.

Smale, M., P.W. Heisey, and H.D. Leathers. "Maize of the Ancestors and Modern Varieties: The Microeconomics of HYV Adoption in Malawi." *Economic Development and Cultural Change*, Vol. 43, No. 2 (January 1995): 351-368.

Souza, E., P.N. Fox, D. Byerlee, and B. Skovmand. "Spring Wheat Diversity in Irrigated Areas of Two Developing Countries." *Crop Science*, (forthcoming).

Byerlee, D. "Modern Varieties, Productivity, and Sustainability: Recent Experiences and Emerging Challenges." In C. Delgado and D. Conyers (eds.), *Post-Green Revolution Agricultural Development Strategies: What Next?*, 1994.

Byerlee, D., K. Khan, and M. Saleem. "Revealing the Rationality of Farmers' Strategies: On-farm Maize Research in the Swat Valley, Northern Pakistan." In R. Tripp (ed.), *Planned Change in Farming Systems: Progress in On-Farm Research* Chichester, UK: John Wiley, 1994.

Fujisaka, S., L. Harrington, and P. Hobbs. "Rice-Wheat Rotation in South Asia: System and Long-Term Priorities Established Through Diagnostic Research." *Agricultural Systems*, Vol. 46, No. 2 (1994): 169-187.

Heisey, P., D. Byerlee, and K. Khan. "Technology Design and System Interactions: Maize Production Systems of the Swat Valley and Mountains." In N.J.R. Allen (ed.), *Himalayan Crucible: North Pakistan in Transition*. USA: St. Martin's Press (forthcoming), 1994.

Morris, M.L., H.J. Dubin, and T. Pokhrel. "Returns to Wheat Breeding Research in Nepal." *Agricultural Economics*, (forthcoming), 1994.

Smale, M. and P.W. Heisey. "Maize Research in Malawi Revisited: An Emerging Success Story?" *Journal of International Development*, Vol. 6, No. 6 (November-December 1994): 689-706.

Traxler, G. and D. Byerlee. "The Economic Returns to Crop Management Research in a Post-Green Revolution Setting." *American Journal of Agricultural Economics*, Vol. 75 (1994): 981-989.

Byerlee, D. "Technology Adaptation and Adoption: Experiences from Seed-Fertilizer Technology and Beyond." *Review of Marketing and Agricultural Economics*, Vol. 61, No. 2 (August 1993): 311-326.

Bohn, A. and D. Byerlee. *The Wheat Breeding Industry in Developing Countries: An Analysis of Investments and Impacts*. Part 1 of *1992-93 CIMMYT World Wheat Facts and Trends*. Singapore: CIMMYT, 1993.

Byerlee, D. "Technology Transfer Systems for Improved Crop Management: Lessons for the Future." In J.R. Anderson (ed.), *Agricultural Technology: Policy Issues for the International Community*. Wallingford, UK: CABI, 1993.

Byerlee, D. "Technical Change and Returns to Wheat Breeding Research in Pakistan's Punjab in the Post-Green Revolution Period." *The Pakistan Development Review*, Vol. 32, No. 1 (1993): 69-86.

Byerlee, D. and T. Husain. "Agricultural Research Strategies for Favoured and Marginal Areas: The Experience of Farming Systems Research in Pakistan." *Experimental Agriculture*, Vol. 29 (1993): 155-171.

Byerlee, D. and M.L. Morris. "Research for Marginal Environments: Are We Underinvested?" *Food Policy*, Vol. 18, No. 5 (1993): 74-84.

Heisey. P.W., K.A. Tetlay, Z. Ahmad, and M. Ahmad. "Varietal Change in Post-Green Revolution Agriculture: Empirical Evidence for Wheat in Pakistan." *Journal of Agricultural Economics*, Vol. 44, No. 3 (1993): 428-442.

Renkow, M. "Differential Technology Adoption and Income Distribution in Pakistan: Implications for Research Resource Allocation." *American Journal of Agricultural Economics*, Vol. 75, No. 1 (1993): 33-43.

Smale, M. and P.W. Heisey. "Simultaneous Estimation of Seed-Fertilizer Adoption Decisions." *Technological Forecasting and Social Change*, Vol. 43 (1993): 353-368.

Negatu, W., W. Mwangi, and T. Tesemma. "Farmers' Varietal Preferences for Durum Wheat in Ada, Lume, and Gimbichu Woredas of Ethiopia." *Ethiopian Journal of Agricultural Science*, Vol. 13 (forthcoming), 1993.

Maredia, M.K., C.K. Eichen, and D. Byerlee. "Strategies for Increasing the Efficiency of National Wheat Research Programs in Developing Countries," paper presented at the XXII International Conference of Agricultural Economists, Harare, Zimbabwe, November 1993.

Hussain, S., D. Byerlee, and P. Heisey. "Impacts of the Training and Visit Extension System on Farmers' Knowledge and Adoption of Technology: Evidence from Pakistan." *Agricultural Economics*, (forthcoming), 1993.

Brennan, J.P., E.J. Warham, and D. Byerlee. "Evaluating the Economic Impact of Quality-Reducing, Seed-Borne Diseases: Lessons from Karnal Bunt of Wheat." *Agricultural Economics*, Vol. 6 (1992): 345-352.

Beyene, H., S. Franzel, and W.M. Mwangi. "Constraints to Increasing Wheat Production in the Small-Holder Sector." In S. Franzel and H. van Houten (eds.), *Research with Farmers: Lessons from Ethiopia*. Melksham, UK:CAB International for the Institute of Agricultural Research, 1992.

Harrington, L.W. "Measuring Sustainability: Issues and Alternatives." In W. Hiemstra, C. Reijntjes, and E. van der Werf (eds.), *Let Farmers Judge: Experiences in Assessing the Sustainability of Agriculture*. ILEIA Readings in Sustainable Agriculture. Chippenham, Wiltshire, UK: Intermediate Technology Publications, 1992.

Dadi, L., W.M. Mwangi, and S. Franzel. "Socioeconomic Constraints to Increasing Maize Production." In S. Franzel and H. van Houten (eds.), *Research with Farmers: Lessons from Ethiopia*. Melksham, UK: CABI International for the Institute of Agricultural Research, 1992.

Lopez-Pereira, M.A., T.G. Baker, J.H. Sanders, and D.H. Meckenstock. "Farming Systems and Adoption of New Agricultural Technologies: An Economic Evaluation of New Sorghum Cultivars in Southern Honduras." *Journal of International Agriculture*, Vol. 31, No. 1 (1992): 55-73.

Brennan, J.P. and D. Byerlee. "The Rate of Crop Varietal Replacement on Farms: Measures and Empirical Results for Wheat." *Plant Varieties and Seeds*, Vol. 4 (1991): 99-106.

Beyene, H., W.M. Mwangi, and W. Negatu. "Research Conducted on Wheat Production Constraints in Ethiopia." In H.G. Mariam, D.G. Tanner, and M. Hulluka (eds.), *Wheat Research in Ethiopia: A Historical Perspective*. Addis Ababa, Ethiopia: Institute for Agricultural Research and CIMMYT, 1991.

Heisey, P.W. and J.P. Brennan. "An Analytical Model of Farmers' Demand for Replacement Seed." *American Journal of Agricultural Economics*, Vol. 73, No. 4 (1991): 1044-1052.

Leathers, H.D. and M. Smale. "A Bayesian Approach to Explaining Sequential Adoption of Components of a Technological Package." *American Journal of Agricultural Economics*, Vol. 73, No. 3 (1991): 734-742.

Low, A. and S.R. Waddington. "Farming Systems Adaptive Research: Achievements and Prospects in Southern Africa." *Experimental Agriculture*, Vol. 27 (1991): 115-125.

Traxler, G., M. Renkow, and L.W. Harrington. "Assessing the Impact of New Technology: Three Levels of Analysis." *Journal of the Asian Farming Systems Association*, Vol. 1 (1991): 227-244.

Heisey, P. "Comment: Maize Research in Malawi." *Journal of International Development*, Vol. 2, No. 2 (1990): 243-253.

Tripp, R. "Does Nutrition Have a Place in Agricultural Research?" *Food Policy*, (December 1990):467-478.

Hassan, R., B. D'Silva, and A. Hallam. "Normative Supply Response Analysis Under Production Uncertainty: Irrigated Multicrop Farming Sector of Sudan." In B. Greenshields and M. Bellamy (eds.), *Government Intervention in Agriculture: Cause and Effect*. International Association of Agricultural Economists, 1990.

Byerlee, D., P. Heisey, and P. Hobbs. "Diagnosing Research Priorities for Small Farmers: Experiences From On-Farm Research in Pakistan." *Quarterly Journal of International Agriculture* Vol. 28, No. 3/4 (1989): 254-265.

Byerlee, D., M. Iqbal, and K.S. Fischer. "Quantifying and Valuing the Joint Production of Grain and Fodder from Maize Fields: Evidence from Northern Pakistan." *Experimental Agriculture*, Vol. 25 (1989): 435-455.

Guido, O. "The Role of Adaptive Research in Increasing Agricultural Productivity: The Experience from Haiti." *Canadian Journal of Agricultural Economics*, Vol. 37, No. 4 (1989): 887-897.

ICARDA

Leybourne, M., F. Ghassali, A. Osman, T. Nordblom, and G. Gintzburger. "The Utilization of Fodder Shrubs (*Atriplex* spp., *Salsola vermiculata*) by Agropastoralists in the Northern Syrian Steppe." *Pasture, Forage and Livestock Program Annual Report, 1993*. ICARDA, Aleppo, forthcoming.

Thomson, E.F., F.A. Bahhady, T.L. Nordblom, and H. Harris. "A Model-Farm Approach to Research on Crop-Livestock Integration: III. Benefits of Crop-Livestock Integration and a Critique of the Approach." *Agricultural Systems*, (forthcoming).

Nordblom, T.L. and F. Shomo. "Food and Feed Prospects to 2020 in the West Asia/North Africa Region," paper prepared for the IFPRI workshop on *Sustainable Agricultural Growth in the Major Ecoregions of the Developing World: Prospects to 2020*, held at Airlie House, Warrenton, Virginia, November 7-9, 1994.

Nordblom, T.L., D.J. Pannell, S. Christiansen, N. Nersoyan and F. Bahhady. "From Weed to Wealth? Prospects for Medic Pastures in the Mediterranean Farming System of North-West Syria." *Agricultural Economics*, Vol. 11 (September 1994): 29-42.

Abdelmagid, S.A. "The Economic Impact of Faba Bean Introduction in Smallholdings: A Case Study of the Gezira Scheme, Sudan." in W. Doppler (ed.) *Farming Systems and Resource Economics in the Tropics*, Vol. 13, Wissenschaftsverlag Vauk, Kiel, 1992.

Maerz, U. "Farm Classification and Impact Analysis of Mixed Farming Systems in Northern Syria." In W. Doppler (ed.), *Farming Systems and Resource Economics in the Tropics*, Vol. 7, Wissenschaftsverlag Vauk, Kiel, 1990.

ICRISAT

Bantilan, M.C.S. and P.K. Joshi. "Tracking the Impact of Wilt Resistant Medium Duration Pigeonpeas in the Semi-Arid Tropics of India." paper presented to 39th annual meeting of the Australian Agricultural Economics Association, Perth, February 1995.

Joshi, P.K. and M.C.S. Bantilan. "Benefits from Improved Soil-Water-Nutrient Management Research: The Case of Groundnut Production Technology." paper presented to 39th annual meeting of the Australian Agricultural Economics Association, Perth, February 1995.

Bantilan, M.C.S. and P.K. Joshi. *Evaluating ICRISAT Research Impact*. ICRISAT, Patancheru, 1994.

ICRISAT. "International Workshop on Integration of Research Evaluation Efforts of ICRISAT with NARS and other International Research Institutions: Program and Participants." ICRISAT, Patancheru, December 1994.

ICRISAT. "Research Evaluation and Impact Assessment for ICRISAT, Annual Report 1993." Resource Management Program - Economics, Patancheru: ICRISAT, 1993.

ICRISAT. *Medium Term Plan 1994-98. Volume I; Main Report*. Patancheru: ICRISAT, 1993.

IFPRI

Rosegrant, M.W. and R.E. Evenson. "Total Factor Productivity and Sources of Long-Term Growth in Indian Agriculture." EPTD Discussion Paper No. 7, Washington D.C.: IFPRI, April 1995.

Craig, B.J., P.G. Pardey, and J. Roseboom. "International Productivity Patterns: Accounting for Input Quality, Infrastructure and Research." Washington, D.C., IFPRI, mimeo, February 1995.

Alston, J.M. and P.G. Pardey. *Making Science Pay: The Economics of Agricultural R&D Policy* Washington, D.C.: American Enterprise Institute Press, forthcoming 1995.

Alston, J.M. and P.G. Pardey. "Revitalizing R&D." in D. Sumner (ed.) *Future Directions in Agricultural Policy*. Washington, D.C.: American Enterprise Institute Press, forthcoming, 1995.

Alston, J.M. and P.G. Pardey. "Agricultural R&D in the Public Interest," report prepared for U.S. Congress, Office of Technology Assessment, forthcoming, 1995.

Alston, J.M., J.R. Anderson, and P.G. Pardey. "Perceived Productivity, Forgone Future Farm Fruitfulness, and Rural Resource Rationalization." Invited plenary paper for theme VII *National and International Research and Technology Transfer*, Twelfth International Conference of Agricultural Economists, Harare, August 1994.

ILRI

International Livestock Centre for Africa (ILCA). *Annual Programme Report 1994*. ILCA, Addis Ababa, Ethiopia, 1995.

International Livestock Centre for Africa (ILCA). *Highlights of ILCA's Programme Achievements, 1986-1991*. ILCA, Addis Ababa, Ethiopia, 1992.

International Livestock Centre for Africa (ILCA). *Improving Livestock Production in Africa: Evolution of ILCA'S Programme, 1974-94*. ILCA, Addis Ababa, Ethiopia. 1994.

International Livestock Centre for Africa (ILCA). *Potential for Impact: ILCA Looks to the Future*. Working Paper 2, ILCA, Addis Ababa, Ethiopia, 1992.

Saleem, M.M.A. and R. von Kaufmann. "Fodder Bank: An Innovation for Improving the Dry Season Nutrition of Traditionally Managed Cattle in the Subhumid Zone of West Africa." ILCA, Addis Ababa, Ethiopia, 1991 (mimeo).

Von Kaufmann, R. and M.M.A. Saleem. "Interactions Between Agronomy and Economics in Forage Legume Research." Addis Ababa, Ethiopia. ILCA Bulletin, No. 35 (1989): 22-27.

IRRI

Antle, J.M. and P. Pingali. "Pesticides, Productivity, and Farmer Health: A Philippine Case Study." *American Journal of Agricultural Economics* Vol. 76, No. 3 (August 1994): 418-430.

David, C.C and K. Otsuka, eds. *Modern Rice Technology and Income Distribution in Asia*. Boulder: Lynne Rienner Publishers, 1994.

ISNAR

Alston, J.M., G.W. Norton, and P.G. Pardey. *Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*. Ithaca: Cornell University Press, 1995.

Various authors. *Training in Planning, Monitoring and Evaluation for Agricultural Research Management*. Series of four training manuals available in English and Spanish. ISNAR and CIAT. 1995.

Zhang, B., P.G. Pardey, S. Fan, and J.M. Alston. "Economics of Technical Change and Productivity Growth in Indonesian Rice and Soybean Production." under review by *Journal of Development Economics*, January 1994.

Craig, B.J., P.G. Pardey, and J. Roseboom. "International Agricultural Productivity Patterns." Center for International Food and Agricultural Policy, Working Paper WP94-1, St. Paul: University of Minnesota, February 1994.

Alston, J.M., P.G. Pardey, and H.O. Carter, (eds). *Valuing UC Agricultural Research and Extension*. Davis: University of California, Agricultural Issues Center, March 1994.

Pardey, P.G. "Economic Perspective on Setting Research Priorities at the Regional Level," paper presented at ISNAR Roundtable on Regional Priority Setting, The Hague: International Service for National Agricultural Research, April 18-20, 1994.

Pardey, P.G. and S. Wood. "Targeting Research by Agricultural Environments." chapter 31 in J.R. Anderson, ed. *Agricultural Technology Policy Issues for the International Community*, Wallingford U.K.: CAB International, 1994.

Peterson, W. "Assessing Organizational Performance: Indicators and Procedures." ISNAR Discussion Paper No. 93-18, rev. 2. The Hague: ISNAR, 1994.

Horton, D. "Assessing ISNAR's Impact." Internal IPR paper, The Hague: ISNAR, 1994.

Horton, D., P. Ballantyne, W. Peterson, B. Uribe, D. Gapasin, and K. Sheridan (eds). *Monitoring and Evaluating Agricultural Research*. Wallingford U.K.: CAB International, 1993.

Wood, S. and P.G. Pardey. "Agroecological Dimensions of Evaluating and Prioritizing Research from a Regional Perspective: Latin America and the Caribbean." ISNAR Discussion Paper No. 93-15. The Hague: International Service for National Agricultural Research, June 1993.

Norton, G.W., P.G. Pardey, and J.M. Alston. "Economic Issues in Agricultural Research Priority Setting." *American Journal of Agricultural Economics* Vol. 74, No. 5. (December 1992): 1089-94.

Fan, S. and P.G. Pardey. *Agricultural Research in China: Its Institutional Development and Impact.* The Hague: ISNAR, August 1992.

Pardey, P.G. et al. *The Economic Returns to Indonesian Rice and Soybean Research.* An AARD/ISNAR report prepared for USAID, Jakarta. The Hague, November 1992.

Horton, D. "Two Institutional Aspects of Impact Assessment: Evaluation of, and in, NARS." In D. Lee, S. Kearl and N. Uphoff. *Assessing the Impact of International Agricultural Research for Sustainable Development*. Ithaca: CIIFAD, 1992.

Alston, J.M. "Research Benefits in a Multimarket Setting: A Review." *Review of Marketing and Agricultural Economics* Vol. 59 (1991): 23-52.

ISNAR. "ISNAR Collaboration with National Agricultural Research Systems: Achievements, Impact and Constraints." Paper prepared for the External Review Panel. The Hague: ISNAR, 1991.

Nielson, J. "Suggested Approaches to Analyzing ISNAR's Impacts." Internal document. The Hague: ISNAR, 1989.