



Expanding the use of impact assessment and evaluation in agricultural research and development

Ronald Mackay^{a,*}, Douglas Horton^b

^a*Concordia University, Department of Education, EN-409, 1455 de Maisonneuve Boulevard West,
Montreal, Quebec, Canada H3G 1M8*

^b*ISNAR, PO Box 93375, 2509 AJ The Hague, The Netherlands*

Abstract

The rationale behind conducting impact assessment studies is the expectation that the findings will be used to bring about improvements in policies and programs and thereby contribute to economic and social betterment. This paper addresses a number of issues related to why impact assessments and other kinds of evaluations are not always utilized in this way. Any particular impact assessment falls along a continuum between academic research at one extreme and utilization-focused evaluation at the other. We argue that impact assessment studies may enjoy greater use by decision makers if they are managed as evaluations conducted towards the utilization-focused end of the continuum. Several factors that influence the use of evaluations are discussed. These include the type of potential use and intended users; the nature of decision-making; factors over and above information that affect how decisions are made; and the various decision levels where evaluations may be used. Direct, indirect and symbolic uses of evaluations are also discussed. Six trends within the professional evaluation community that tend to enhance the use of evaluation are described. Finally, we offer some practical suggestions for enhancing the usefulness and use of impact assessments and other evaluations within the agricultural research and development community.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Evaluation; Impact assessment; Agricultural research; International development; Utilization-focused evaluation

* Corresponding author.

E-mail addresses: ronald.mackay@sympatico.ca (R. Mackay), d.horton@cgiar.org (D. Horton).

1. Introduction

In this paper we address the concern expressed by organizations in the international agricultural research community (Pingali, 2001; Horton, 1986, 1998), that impact assessment studies have made less of a difference than might be expected and much less than desired by those who conduct them. Undertaking impact assessment studies is not enough. “The information obtained from them has to be utilized by the stakeholders involved” (Rutherford et al., 2001: 45). We address the issue of utilization from the perspective of the professional evaluation community, which concerns itself with impact assessment along with other types of evaluation. Professional evaluators have already faced—and boast some success in addressing—the challenges of under-utilization of evaluation studies.

After this Introduction, in Section 2, we discuss some distinctions between research and evaluation and present the notion of a research-evaluation continuum. We argue that to enhance use, impact assessment should be managed as a form of evaluation rather than as a form of research. In Section 3, we discuss the potential uses of impact assessment and of evaluation more broadly and identify a number of interrelated factors that influence how information is used in decision making. In light of a discussion of these factors, we suggest that impact assessors in the international research and development community may be looking for direct use of their findings at levels where they are least likely to occur—the levels of strategy and policy. By the same token, they may be failing to cultivate use at the operational levels, where it is more likely to occur. In Section 4, we discuss six trends that reflect how the professional evaluation community has responded, with some success, to the challenge of promoting evaluation use. Section 5 suggests some practical ideas that evaluators might consider to promote the use of impact assessment and evaluation in research and development organizations.

In this paper, evaluation in the CGIAR (Consultative Group on International Agricultural Research) is used to illustrate many of the ideas discussed. The CGIAR, created in 1971, is a group of 16 centres that work in more than 100 countries to mobilize cutting-edge science to reduce hunger and poverty, improve human nutrition and health, and protect the environment. Readers are invited to reflect on these ideas and their implication for the organizations within which they work, or with which they are associated.

2. The research — evaluation continuum

Over the past 30 years, and particularly during the last decade, evaluation has developed rapidly as a professional field. The origins and development of agricultural research evaluation have been documented by Horton (1998). Evaluation textbooks generally distinguish research from evaluation (e.g. Worthen et al., 1997; Patton, 1997; Weiss, 1998). Research is viewed as systematic investigation in a particular field of knowledge that is carried out in order to establish laws or principles. Scientific research typically involves an experimental design capable of discerning

causal relationships, the collection of objective, reliable and valid quantitative data, and the use of statistical analytical techniques. Research tends to be motivated by the curiosity of the inquirer. It seeks a scientific explanation for relationships between variables to draw conclusions and arrive at generalizations, hypotheses, principles and laws. Research is typically judged by technical standards of reliability and validity defined by the discipline in which it is carried out. Its results tend to be directed at the researchers' disciplinary peers. Researchers acquire their training within their substantive discipline (Horton, 1998).

In contrast, evaluations are carried out to judge the worth, merit or quality of some object (termed the evaluand)—whether it is proposed, ongoing, or completed—in terms of its relevance, effectiveness, efficiency and impact. Evaluation is typically motivated by the need to resolve practical policy and program related issues. It tries to explain how specific policies and programs bring about their results within a given context. Evaluation information is generated for and directed at decision makers and time is of the essence. Particular evaluations are judged by the usefulness of their findings, the practicality of their methods, their respect for propriety and the accuracy of the data and results. Evaluation questions tend to cross disciplinary boundaries and so demand mixing methods rather than disciplinary orthodoxy.

The distinctions just made between research and evaluation are, in fact, relative rather than absolute, and for this reason it is useful to view “pure research” and “pure evaluation” as the extreme points of a continuum. Following this logic, specific research and evaluation studies can be located at different points along the continuum.

2.1. Research-oriented evaluation

Research-oriented evaluation or “evaluation research” has been defined as the systematic application of social science research procedures in assessing the design, implementation and utility of policies and programs. At this extreme the principal, and possibly the only, factor that distinguishes evaluation research from academic research is the focus of the former on issues relevant to current policy. Evaluators working at the research-oriented end of the continuum consider that “evaluation may be successful even if the information provided is ignored, or even misused. Once the findings are presented in a clear and accessible fashion, the evaluation is over. What follows is certainly critical, but is essentially a political process. Interested evaluators are best off observing the action at some distance, preferably through heavy lenses” (Rossi, 1987: 10). This extreme view is hardly helpful for overcoming the limited use of impact assessments and their limited impact on donors' willingness to fund research and development projects and programs (Bennett, 2001; Pingali, 2001; Ndiritu, 2000; Maredia et al., 2000).

2.2. Utilization-focused evaluation

At the other end of the continuum, utilization-focused evaluation is described as “evaluation done for and with specific, intended primary users for specific, intended

uses” (Patton, 1997: 23). The practice of utilization-focused evaluation does not limit itself to the application of methods prescribed by any one discipline. It emphasizes systematic data collection but chooses the methods it employs so as to address the information needs of the intended user. Hence, methods may vary from evaluation to evaluation. Evaluation practiced from this perspective is described as the “systematic gathering, analyzing and reporting of information about a policy, program, service, or intervention for use in making decisions” (Hudson et al., 1992: 2). The focus is on *use to decision makers*. This focus introduces what the evaluation community refers to as the *personal factor*. Patton (1997), a leading exponent of utilization-focused evaluation, argues that evaluation studies should target specific users (clearly identified people), not faceless stakeholder groups or organizations.

2.3. *Implications for impact assessment in research and development organizations*

The conclusion to be drawn from this brief discussion of research-oriented and utilization-focused evaluation is that, if evaluations are to be useful and influential, choices about the approach and methods should reflect the information needs and circumstances of identified, intended users, not the disciplinary concerns of the evaluator. If an impact assessment is conducted as a research study addressed primarily to the researcher’s disciplinary peers, it is unlikely to be found useful by those charged with making policy and other decisions. In contrast, an impact assessment that is designed and conducted to meet the specific information needs of specific decision makers is more likely to be of use to, and used by, them.

3. Factors influencing the use of evaluations

Information use is often thought of in binary terms—“use” and “non-use”. In fact, research shows that processes of information use are highly complex. In this section we address some of the factors involved. These include: the types of *potential information users*, the decision making *processes*, the multiple *factors* that affect decisions (in addition to information), the differing *types of use* of evaluation and the *decision-making level* at which evaluative information may be used. We close by noting that involving stakeholders in evaluation processes may in some cases have as great an influence on decisions as the dissemination of evaluation findings in formal reports.

3.1. *Potential or intended users*

It is a commonplace to say that there are multiple audiences for impact studies. A more precise identification of the intended user of a given set of findings, however, is of paramount importance. Without knowing the intended user, it may not be possible to communicate the results of an impact assessment in a form that the target

audience can readily understand and via a channel that they can readily access. Professional journals may be suited to one audience but miss other important potential users.

The papers presented at a meeting convened in Rome in May 2000 to consider the future of impact assessment in the CGIAR (TAC Secretariat, 2001) refer to a wide range of potential users of impact assessment, including donors and development agencies, policymakers, managers and researchers of international and national centres, scientific and professional peers, philanthropic organizations, taxpayers, and consumers. However, the “personal factor”, referred to above, insists that we must target specific persons not abstract institutions, organizations or categories of stakeholders. This means that the evaluator must make a concerted effort to locate specific people who need specific information and who are also able and willing to use it when provided. Targeting individuals, of course, does not necessarily commit them to using the findings neither in the way nor to the extent that the evaluator would ideally like. It does, however, ensure the relevance and accessibility of evaluation information to the intended user.

Moreover, as those active in the international research and development community observe (Kumar, 1995; Bennett, *op. cit.*; Ndiritu, *op. cit.*), the potential for utilization will be enhanced not only if the study is conducted *for* identified stakeholders but also if the study is carried out *with* the active participation of these identified stakeholders in one or more of its phases. We have more to say about stakeholder participation in evaluation studies in Section 4.

3.2. *Decision-making processes*

Classical decision theory assumes that decisions are made rationally. It assumes that decision makers seek the best possible alternative available in order to maximize the achievement of policies, goals and objectives. The rational decision-making model presents information use as being integrated within a series of sequential steps (Fig. 1).

Research on decision-making indicates that the rational model is unrealistic, assuming information processing capacities and a level of rationality that most policymakers and managers simply do not and cannot possess. Decision-makers tend to be more pragmatic than purely rational and seek, not the ideal or even the best solutions to their problems, but rather “satisfactory” solutions. This is because solutions are often “path dependent”, in the sense that they are constrained by previous decisions. This more common kind of pragmatic decision-making behaviour is referred to, without any pejorative connotations, as “sufficing”. This broader incremental strategy is widely practiced by decision makers, especially when issues are complex, uncertain and have the potential to engender conflict.

Most decision-making, following the incremental model, involves accretion-like, risk-reducing processes of “muddling through” (Lindblom, 1980). In such a context, evaluation findings are seldom used in the ways or to the extent that evaluators would like them to be.

In the rational model, the decision maker:	In the incremental model, the decision maker:
Acknowledges a problem	Acknowledges a problem
Develops goals and objectives for its resolution	Limits the number of alternatives considered
Generates all of the possible alternative solutions	Considers only those that closely approximate to the current situation (major change is eliminated)
Considers the consequences of each alternative	Reduces the complexity of decision-making so as to be manageable
Reviews all the alternatives in terms of their capacity to meet the goals and objectives	Makes the most of experience and minimizes the need for new knowledge
Selects the best alternative – the one that maximizes the achievement of the goals and objectives	Increases the accuracy of predicted consequences
Implements the decision and monitors the implementation process	Conserves time and energy
	Promotes modest courses of action
	Minimizes risk

Based on Hoy and Miskel, 1996 pp. 267 and 280.

Fig. 1. Rational and incremental models of decision-making.

3.3. Factors that influence decision-making

The political scientist and evaluator Carol Weiss (1999: 477) has identified four key sets of factors that influence decisions, which she refers to as “the four Is”—*interests, ideologies, institutions* and *information*. Decision makers tend to seek advantage for their own causes and constituencies. Hence personal *interests* can have a strong influence on decisions. In many instances, executives, board members, program staff or others attempt to put aside evaluation results that are not in accord with their personal interests.

Ideologies are also powerful forces guiding policy makers. Most democracies are made up of political parties that represent more or less distinct ideological positions. Policymakers are expected to act within the ideologies embraced by their parties. Similarly, research organizations have their ideologies. If, for example, an evaluation were to conclude that the goals of the CGIAR could be achieved more efficiently if national research organizations were bypassed in favour of targeting national politicians and producers directly, it is unlikely that this would affect decision-making in the system. The strong CGIAR ideological principle of partnerships with national agricultural research organizations would prevail.

The history and culture of *institutions* exert a powerful influence on what and how decisions are made. Decisions are never made in a vacuum, and are often dependent upon how successions of previous decisions were made. The general direction has usually already been established or constrained by other factors and so fundamental changes in direction, based solely on information may not be possible. To continue

with the CGIAR as an example, even though an evaluation at the system level of the CGIAR might conclude that it would be most effective to merge two or more of the 16 independent centres, such a recommendation may not be acted on because of effective opposition by the concerned centres, their boards or their host country governments.

The fourth “I”, *information*, is of course also important, but one must keep in mind that evaluation findings are only one source of information among many others that are constantly pouring into decision-making processes. Interest groups, advisors, think tanks and many others may also be providing information and opinion to decision makers, and these sources may carry more weight than findings from an evaluation study.

3.4. *The 4 Is’ in operation—a concrete example*

Over the years, there have been many studies by ISNAR, the only CGIAR Centre whose primary mandate is to strengthen the management capacity of NARS, that have pointed out the problems of rate-of-return studies and the limitations of mainstream economic impact assessments as a means of promoting organizational learning. Relevant studies include Horton (1986,1998), Goldsmith (1993), Horton and Mackay (1998), Horton et al. (1993), Mackay et al. (2002). However, these studies and their findings appear not to have been taken on board.¹ Weiss’ “four Is”—ideology, institutions, interests, and information—provide a framework that helps explain why these arguments have been, by and large, ignored.

3.4.1. *Ideology*

The CGIAR has a deeply rooted “hard-science” culture and an unquestioning belief in the value of applying modern, positivist scientific methods to the solution of agricultural, environmental and related problems. In this context where science is virtually synonymous with quantitative analysis the alternative approaches we have been advocating, which include qualitative methods, participatory approaches and constructivist logic, are regarded as highly suspect—‘soft science’ at best.

A second aspect of CGIAR ideology is an insufficiently critical view of the value of the work of the CGIAR centres themselves, which are often referred to as “Centres of Excellence”. A review of the CGIAR system conducted in the mid-1990s (CGIAR Secretariat, 1998: 6) made the following assessment:

The CGIAR has established a universally acknowledged record of success in international agricultural research. Indeed, investment in the CGIAR has been the single most effective use of official development assistance, bar none. There can be no long-term agenda for eradicating poverty, ending hunger, and ensuring sustainable food security without the CGIAR.

This view is echoed in numerous CGIAR-related documents and supported by numerous economic impact assessments commissioned by CGIAR centres and

¹ Stephen Biggs, personal communication, 8 January 2003.

donors. The strong belief in the central value of bio-physical research and the excellence of the work of CGIAR centres, which has prevailed within the system over the years, discourages openness to learning and to broader types of evaluation, which question basic assumptions and premises and seek to learn from errors and failures as well as from successes.

This tendency to cling to a single disciplinary perspective can be explained in terms of what is currently known about organizational learning. Organizational learning does not occur automatically when new information becomes available, irrespective of how well-founded the new information may be. For new information to be shared throughout a community of practice it must be given a commonly understood meaning by the concerned individuals and groups. All communities of practice possess prior mental models which influence their interpretation and formation of new information. Before new information can “make sense” it is compared with existing belief structures and found to be either compatible or incompatible. Compatibility between existing mental models and new information is essential for individual, group and organizational learning to occur. Individuals, groups and organizations learn only to the extent that their existing belief structures are open to admit new information and permit varied interpretations. If the mental models are so rigid that they are impervious to change, little or no new learning can occur (Huber, 1991; Dodgson, 1993).

3.4.2. *Institutions*

From the beginning, impact assessment and evaluation in the CGIAR have been associated with accountability and public awareness, and these links have been institutionalized. External reviews are a key component of the CGIAR’s accountability system (Özgediz, 1999) and the main focus of the Standing Committee on External Reviews (EPMRs) is accountability, not learning or change. Similarly, the Standing Panel on Impact Assessment (SPIA) is charged with ex post impact assessment, primarily to meet donor needs related to accountability.² The establishment of SPIA (formerly the Impact Assessment and Evaluation Group) in the mid-1990s grew out of efforts of the CGIAR’s Public Awareness and Resource Mobilization Committee to obtain better evidence of the value of the CGIAR’s work for use in public awareness activities. As a result, neither of the two institutional bodies responsible for system-wide evaluation and impact assessment is concerned formally with the use of evaluation to promote institutional learning and change within the CGIAR system.

3.4.3. *Interests*

The third group of factors that discourages the introduction of alternative evaluation approaches in the CGAIR relates to the personal and professional interests of key stakeholders. Over the years, most of those outside the CGIAR system who have pressed for impact assessment and evaluation have been donors who sought information to satisfy their own accountability needs, and these groups promoted

² Personal communication from Hans Gregersen, Chair of SPIA, 2 April 2002.

the use of economic impact assessment and expert reviews (the traditional method used in EPMRs). More recently, some donors are promoting the application of indicator-based performance measurement systems. Within the CGIAR, impact assessment and evaluation activities have generally been assigned to agricultural economists, who for obvious professional reasons, have tended to perpetuate the use of economic frameworks and quantitative methods. Key positions, both in individual centres and in system-level bodies (such as the Technical Advisory Committee, the interim Science Council, the CGIAR Secretariat, and the Executive Council) have been occupied by economists, who have fostered the use of economic methods in planning, priority setting and evaluation.

3.4.4. Information

Finally, decision-makers in the CGIAR and in the donor community are influenced by many sources of information. It is unlikely that the studies we have published in professional journals influenced many of these individuals. Indeed our publications are unlikely to have reached most decision makers. This underlines the limited role that academic journals, each with its relatively narrow professional readership, may play in influencing policy and management decisions.

3.4.5. Prospects for the future

A number of current trends may be creating more openness to innovative evaluation approaches in the CGIAR. The accelerating processes of social, economic, political and institutional change and the emerging new modes of production of knowledge and technology (Gibbons et al., 1994; Gibbons, 1999) are undermining traditional beliefs in the Newtonian model of science and in the CGIAR centre as an ideal structure for organizing international agricultural research and development activities. As a consequence, there are moves, both within CGIAR centres and in the donor community, to foster new forms of evaluation that promote institutional learning and change. In recent years, a new generation of researchers has also entered the CGIAR with training and expertise in such fields as sociology, ecology, science and policy studies, and political science. These new professionals have begun to experiment with new types of evaluation, as evidenced in the other papers in this issue of *Agricultural Systems*.

At the same time, as the CGIAR system finds it more and more difficult to fund its programs and is subject to expanding demands for accountability, the challenge will be to develop *both* types of evaluation and balance the demands for accountability against the need for evaluations that foster institutional learning and change.

3.5. Uses of evaluation findings

We typically think in terms of the direct influence of evaluations on decisions, but evaluation findings are used in many other ways as well. Weiss (1998) and others have identified three basic types of use of evaluation findings: direct, indirect and symbolic. Direct or instrumental use occurs when information or findings are

applied directly to change an action or alter a decision. Indirect or conceptual use refers to a more intellectual and gradual process in which the decision maker is led to a more adequate appreciation of the problems addressed by the policy or program. Symbolic use refers to situations where the evaluation results are accepted on paper or in public pronouncements, but go no further. Many evaluations are symbolic in that they are carried out simply to comply with administrative directives or to present an image of “modernity”. Research on evaluation use indicates that symbolic use is quite common. Indirect use, often through processes of what Weiss terms “enlightenment” are also common. Direct use of evaluation findings in decision-making is rather infrequent, especially where politically sensitive issues—such as funding or continuation of a program—are at stake.

3.6. *Decision-making levels*

Evaluation findings can have an influence at different decision-making levels within organizations and administrative systems. Balthasar and Rieder (2000) define three key levels as the *operational level* of day-to-day activities and procedures, the *structural level* where organization-wide decisions are made and organizational arrangements are determined, and a higher *strategic level* where overall policy decisions are taken.

Research on evaluation use and organizational learning carried out by these authors suggests that the main direct use of evaluation findings is at the operational level. Uses of findings at the structural and strategic levels are generally indirect (Fig. 2).

Direct use at the operational level can be promoted by involving program staff and other interested parties in planning and conducting the evaluation. Promotion of indirect use of evaluation findings at the structural and strategic levels is a much more complex process, which requires keeping in touch with decision makers at these levels, and exploiting “windows of opportunity” when they occur. Such opportunities may occur during an evaluation or long after it has been completed.

3.7. *Process use*

Evaluators generally think in terms of enhancing the use of the evaluation *findings* presented in their reports. But recent studies and experiences indicate that evaluation *processes* can also have strong influences on decisions and actions. This phenomenon has been defined as “process use” (Patton, 1997: 87–113). A growing body of evidence and experience indicates that participation in one or more of the phases of evaluation—planning, design, data gathering, analysis, interpretation, and write-up—can have significant and lasting effects on the knowledge, attitudes and skills of people and on their actions. Horton et al. (2000) and Mackay et al. (2002) refer to the advantages of process use observed in self-assessment exercises undertaken by managers and staff members in several agricultural organizations in Cuba, Venezuela and Panama.

Decision-making level:	Type of use:		
	Direct / Instrumental	Indirect / Conceptual	Symbolic
Strategic		✓	✓
Structural		✓	✓
Operational	✓		✓

Based on Balhassar and Rieder, 2000, Table 1.

Fig. 2. Principal types of evaluation use at strategic, structural and operational levels.

3.8. Summary

Many factors influence the use of evaluations in decision-making. Evaluation use cannot be reduced to the simple dichotomy “use” or “non-use”. Information from an impact assessment may serve to enlighten those charged with policy and strategic decisions in ways that are difficult to predict or detect. On the other hand, evaluation results may have a more direct use at the operational level. Involving program staff in an evaluation may also have a significant influence on future decisions—in many cases such “process use” may be more important than the use of findings distributed in a formal evaluation report.

4. Trends within the evaluation community

In this section, we briefly describe six on-going trends within the evaluation community that tend to promote the use of evaluations.

4.1. Professionalization of evaluation

For over two decades, there has been a strong movement to professionalize evaluation. Numerous professional evaluation associations have been established. The largest is the American Evaluation Association (AEA) which currently has more than 3000 members from 54 countries. The Canadian Evaluation Society has some 1500 members. Regional and national evaluation societies have also been established throughout the world.³

A primary role for professional evaluation associations is to promote good practice by establishing standards, principles and guidelines for evaluation practice. The most thoroughly elaborated and influential standards are the Program Evaluation Standards endorsed by the AEA (Joint Committee, 1994). They have been refined

³ Information on the professional evaluation organizations and related information can be found on the websites of the American Evaluation Association (www.eval.org), the International Organization for Cooperation in Evaluation (home.wmis.net/~russon/ioce/) and M and E News (www.mande.co.uk/news.htm).

over a 25-year period. The standards were prepared for the evaluation of education and training programs, projects and materials, but they have broad applicability to evaluation efforts in other sectors. Hence, they have formed the basis for standards and complementary principles adopted by other evaluation associations.

The Program Evaluation Standards include seven *utility* standards that are intended to ensure that an evaluation will serve the information needs of intended users. Three *feasibility* standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic and frugal. Eight *propriety* standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results. Twelve *accuracy* standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

The order in which the four groups of standards are listed is significant. The 12 groups which sponsored the standards felt that evaluators were placing too much emphasis on measurement and technical accuracy at the expense of propriety, feasibility and especially utility. The standards make a point of emphasizing that an evaluation study that is technically excellent but is not useful, is deficient. The information generated by an evaluation must be useful for the intended users.

Professional evaluation associations also seek to enhance the professional development of their members with a view to establishing standards of expertise for possible future professional certification⁴—in much the same way as other professions grant accreditation. Numerous universities now offer courses in evaluation and some grant advanced degrees in this field. There has also been a proliferation of professional journals and books. Evaluation is emerging as a “transdiscipline” (Scriven, 1993; Worthen et al., 1997) and many professionals with academic roots in various disciplines now consider themselves to be *evaluators*.

To sum up, the field of evaluation is rapidly growing in ways that strengthen its professionalism, and one of the profession’s main concerns is enhancement of the use of evaluation results.

4.2. *Expansion of formative, participatory, empowerment and developmental evaluation*

The second trend is the expansion of formative evaluation, including its participatory, empowerment and developmental branches. Until the late 1950s, evaluation in the public sector used mainly experimental or quasi-experimental designs to focus on public programs for accountability purposes. It was expected that results of these rigorous studies would lead to decisions for continuing or eliminating various programs (Russ-Eft and Preskill, 2001: 43). This expectation was often frustrated, as results proved inconclusive and decisions continued to be influenced more by political considerations

⁴ A report (Long and Kishchuk, 1997) commissioned by the Canadian Evaluation Society on the pros and cons of professional certification for evaluators can be found on the CES website (<http://www.evaluationcanada.ca>).

than by evaluation findings. Some evaluators began to question the utility of large-scale impact studies and pointed out that evaluations could also provide useful information for improving on-going programs.

The terms *formative* and *summative* as applied to evaluation were coined three decades ago (Scriven, 1967) to distinguish between two of evaluation's main purposes or functions. *Formative evaluation* is aimed at providing information to the staff of a program on how to improve the program and its results. It is typically undertaken during the implementation phase of a program. In contrast, *summative evaluation* produces a summary statement about a program's effectiveness, value, or impact. A summative evaluation is conducted after completion of a new program, or after stabilization of an on-going program, for the benefit of an external audience or decision maker. A recent survey of AEA members confirms that formative evaluation is much more prevalent today than summative evaluation (Russ-Eft and Preskill, 2001: 21–22).

Formative evaluation is often carried out with the direct involvement of those responsible for the program being evaluated. Involving program staff and stakeholders in an evaluation can improve the quantity and quality of information gathered as well as improve the analysis and interpretation of results. Involvement of program staff also contributes significantly to the understanding and acceptance of evaluation results and their use in improving programs and their results. Thus, formative evaluation is increasingly used as a training ground to build the capacities of individuals and organizations to employ evaluation logic and the discipline of evaluative thinking in the service of learning how the policy or program they are responsible for actually works and how it can be improved (Picciotto, 1995; Scriven, 1998; Horton et al., 2000).

Formative evaluation has evolved into several different practices. These include “participatory evaluation” (Whitmore, 1998), “empowerment evaluation” (Fetterman, 2001) and “developmental evaluation” (Patton, 1997). While each has its own characteristics, they all usually involve some form of partnering relationships between an evaluator and clients who are engaged in policy or program development over time. In participatory evaluation, empowerment evaluation and developmental evaluation, the interaction and involvement of program staff in evaluative activities may be as important as the preparation of a traditional evaluation report, or even more so.

Formative evaluation and the novel practices mentioned have led to many new uses of and users for evaluation far beyond those present in traditional summative evaluation, carried out to meet the accountability requirements of external stakeholders.

4.3. *Shift from method-oriented to theory-driven evaluation*

The third trend is the shift from method-oriented to theory-driven evaluation. Initially, evaluators typically worked within the scientific discipline they had been trained in, and employed the theories and methods of that discipline to determine the merit and worth of programs. This disciplinary orientation led to a simple input/

output or black box type of evaluation. Chen (1990: 18) summarizes this approach as follows:

“This type of evaluation is characterized by a primary focus on the overall relationship between the inputs and outputs of a program without concern for the transformation processes in the middle. Such simple input-output or black box evaluations may provide a gross assessment of whether or not a program works but fail to identify the underlying causal mechanisms that generate the treatment effects, thus failing to pinpoint the deficiencies of the program for future program improvement or development. A black box evaluation is usually not sensitive to the political and organizational contexts of input and output, and it neglects issues such as the relationship between the delivered treatment and the planned treatment, between official goals and operative goals, or between intended and unintended effects.”

Since the mid-1980s, many evaluators have argued that evaluations should be driven not by the theories and methods of particular scientific disciplines but by the theories and assumptions underlying the programs being evaluated. Program theory refers to the set of beliefs that underlie action. Program theory is “a plausible and sensible model of how a program is supposed to work” (Bickman, 1987: 5). It identifies “program resources, program activities, and intended program outcomes, and specifies a chain of causal assumptions linking program resources, activities, intermediate outcomes, and ultimate goals” (Bickman, 1987: 78).

As a general rule, the policy or program being evaluated will not have a well-elaborated theory, and an important task of the evaluator will be to work with program staff to reconstruct and develop consensus on a program theory that can then help guide the design of the evaluation and the interpretation of results. In many cases, the exercise of developing a policy or program theory helps staff members recognize weaknesses in the design and implementation of the intervention, even without the further collection and analysis of information on the program’s operations and results (Patton, 1997: Chapter 5).

A program theory often takes the form of a flow chart with a series of steps, and accompanying assumptions, leading from resources and activities to outputs, intermediate outcomes and ultimate impacts. The evaluator can organize the evaluation to collect data and test assumptions at various points along the anticipated chain of cause/effect relationships.

The expanding practice of theory-driven evaluation has contributed to the usefulness and use of evaluations for both program staff and external stakeholders who need to understand the programs they support, the results they produce, and how these results are achieved.

4.4. Development of internal evaluation

The fourth trend we identify is the development of internal evaluation. Internal evaluation is “a systematic process for collecting and analyzing reliable performance

data and interpreting the results to inform and clarify issues that will illuminate and facilitate the decision-making process within organizations” (Sonnichesen, 2000: 3). Internal evaluation covers self-evaluation as well as assessments conducted by or with the assistance of external consultants working to terms of reference that originate within the institution or organization.

Experience in a broad range of organizations confirms that those responsible for policy and decision making are more likely to understand and use evaluation results if they have been directly involved in the evaluation process.

There is a strong trend to develop evaluation capacity within organizations. Internal evaluation is growing rapidly in response to disenchantment with external evaluation, funding cuts for large-scale evaluations and widespread concern with the poor utilization of external evaluations. Private firms, government departments and NGOs are aware of the potential value of internal evaluation not only to monitor activities but also to promote system-wide learning and continuous improvement of internal processes, products and services (Love, 1991). Some of the challenges encountered in developing capacity for participative internal evaluation in agricultural research organizations are discussed in Horton et al. (2000) and in Mackay et al. (2002).

Sonnichesen (2000) notes that with the rise of internal evaluation, lengthy program evaluations are being replaced by shorter, management issue-oriented studies. Evaluators are becoming more aware of the information needs of clients, and client focus is driving the internal evaluation process. Analysis, not compliance or accountability, is becoming the focus of evaluation. A growing diversity of evaluation approaches is being applied, as “the extraordinary complexity and ambiguity of many modern organizational problems hinder resolution using a single approach” (Sonnichesen *op. cit.* p.11). Additional tasks, such as planning and performance measurement are also being undertaken by internal evaluators.

4.5. Diversification and mixing of methods

The early method-oriented evaluations generally employed quantitative, experimental or quasi-experimental designs to assess program results. However, as the purposes and users of evaluation broadened and evaluators attempted to answer more questions associated with the design, underlying theories and results of programs, they have found it necessary to employ a wider range of methods to collect and analyze data and interpret findings. Russ-Eft and Preskill (2001: 46) describe this trend:

“Over the last 35 years, evaluation has grown from being monolithic in its definition and methods to being highly pluralistic. It now embraces multiple methods, measures, criteria, perspectives, audiences, and interests. It has shifted to emphasizing mixed method approaches in lieu of only randomized control group designs, and has embraced the notion that few evaluations are value-free and, by their very nature, are politically charged.”

Evaluators must be able to make appropriate, selective use of a wide range of evaluative techniques to assess processes and results at the level of individuals,

organizations and larger systems. The use of multiple methods to collect both quantitative and qualitative information is made necessary because no single data collection method or data set is sufficient to provide an adequate account of the complexities of a program or wider system as it operates in a rapidly changing environment.

In the 1970s and 1980s there was considerable debate within the evaluation community on the validity and utility of quantitative versus qualitative methods. In the 1990s, the so-called “paradigms debate” diminished considerably. The strengths and weaknesses of both quantitative and qualitative methods have been acknowledged. The consensus is that no single method or data set is capable of capturing adequately the complexity of relations among policies, organizations, programs, and people. A more mature view has emerged which emphasizes matching evaluation methods to the questions directing the evaluation and the information required by the intended users. In order to provide intended users of evaluation with adequate information and sound conclusions, leading evaluators now generally recommend the use of mixed method evaluation approaches employing both quantitative and qualitative data (Greene and Caracelli, 1997; Patton, 2002).

Restricting or distorting questions to those which can be answered by an evaluator’s preferred method is not a valid option in today’s utilization-focused evaluation community. The practice of mixing methods has broadened the range of questions and issues that evaluators can systematically address and has improved communication between researchers and intended users.

4.6. *Attention to the “soft side” of evaluation: communication, negotiation and facilitation*

As the questions that evaluators address and the roles they play have diversified, there has been increased interest in the “soft side” of the evaluation enterprise—communication, negotiation, and facilitation of evaluation processes. Inadequate consideration of target audiences, excessive length and complexity of reports, poor timing, obscure technical jargon, and limited distribution of evaluation reports have long been recognized to discourage the utilization of evaluation findings (Worthen et al., 1997; Winberg, 1992; Guba and Lincoln, 1981). Evaluators now emphasize the importance of *two-way communication* between the evaluator and potential users, to ensure mutual understanding throughout the evaluation process, from the formulation of evaluation questions all the way to the design of potential solutions to problems brought to light by the evaluation.

Several measures have been suggested in the evaluation literature to enhance communication in the interest of making evaluation processes and findings more relevant and therefore more likely to be used (Torres et al., 1996). Evaluators are urged to:

- seek opportunities for active, long-term involvement with programs and organizations, to build trust and committed relationships with those who make policy and management decisions;

- work with clients and stakeholders in order to learn who needs what information, when and for what purpose;
- engage in multifaceted forms of inquiry in which any and all data that will shed light on a question are pursued;
- spend time working with clients and stakeholders in order to encourage them to take ownership of the evaluation process and results;
- communicate in a number of ways, including informal one-on-one and group meetings, in addition to formal reports;
- match the format and communication style of their reports to the characteristics and needs of each of their audiences.

5. Suggestions for enhancing the use of impact assessment and evaluation

We end this paper by suggesting a few practical ways in which evaluators might enhance the utilization of impact assessment in the agricultural research development community. In doing so, we continue to use the CGIAR as an example. Many of the characteristics of the CGIAR are shared by other international agencies, in which the same or similar lessons may apply.

5.1. Manage impact assessment as an evaluation activity in support of decision-making

To date, impact assessment has been largely separated from other types of evaluation and only weakly linked to planning and management decision-making in the international agricultural research community. There is also an unfortunate association of impact assessment with public awareness and fund raising. There are few evaluation units within the 16 international centres that comprise the CGIAR. The role of impact assessor or “research evaluator” tends to be adopted by agricultural economists on an “as needed” basis. Consequently, impact assessment tends to be seen as a specialized task of agricultural economists, aimed at producing information to promote public awareness and to satisfy external accountability requirements. It is less often viewed as an ongoing function in support of management and decision-making. Experience in the broader evaluation community indicates that for impact assessment to make more of a difference, it should be managed as a type of evaluation rather than as research. For this to occur, there would need to be a concerted move within the agricultural research community to develop more integrated management systems in which evaluations supported decision-making at various levels (Horton and Dupleich, 2001). The CGIAR’s Standing Panel on Impact Assessment and the Standing Committee on External Reviews might be well placed to foster such an initiative. The recent CGIAR initiative to determine how the poor have benefited from agricultural research may represent a step in this direction (iSC Secretariat, 2002: 9–11).

The idea of having an independent evaluation function—the current SPIA of the CGIAR—within the international agricultural research community dates back less than 10 years. It may be instructive to examine other experiences with the establishment of evaluation units or groups to see what might be learned. One lesson seems to be that unless they rapidly establish their usefulness and credibility within the system they are intended to serve, their survival is far from guaranteed.

Based on experience within the Government of Canada, [Mayne reported \(1992\)](#) that a constructive way to start up an evaluation unit is to clarify and systematically build upon the most useful role that the unit may play within the larger system. An environmental scan can help the unit to situate itself within the system-wide management and decision-making structures. Mayne suggests that the manager and staff of the evaluation unit must, at the outset, seek answers to such questions as: What is the role envisaged for the impact assessment function? How are decisions taken within the larger system? What kinds of program and policy decisions are taken, by whom and when? Who are the key players in the system? What are the kinds of issues they are likely to have to deal with over the next year or two? What kinds of information are they likely to need? With well-informed answers to questions like these, it is possible to sculpt the most useful role or roles that the unit might play within the system. [Chelimsky \(2001\)](#) and [Picciotto \(2002\)](#) discuss similar issues for evaluation units within foundations and the World Bank.

5.2. *Attend to practical use*

Impact assessment in the international agricultural research community is located toward the research end of the research-evaluation continuum. Impact assessors generally prize technical merit over practical utility. Experience has shown that to rely solely upon technical merit to win support for evaluation is unwise ([Mayne, op. cit:](#) 308). [Thoening \(2000: 223\)](#) warns that “there is no surer way of stifling evaluation at birth than to confine it to the ghetto of methodology”. Initially, potential users may give impact assessors the benefit of the doubt. But to prove its worth, over time, impact assessment needs to shift towards the “utilization-focused” end of the research-evaluation continuum. [Patton \(1997:58\)](#) suggests that impact assessment practitioners can build greater concern for practical use into their studies in the following concrete ways:

- identify intended users and intended uses at the outset of the impact assessment study;
- include intended users in the initiation and execution of the evaluation study;
- target specific persons as users, as opposed to abstract institutions, organizations or categories of stakeholders;
- question the assumption that the evaluation’s funder is automatically the primary stakeholder;
- overcome the tendency to remain above or outside the fray of people and politics.

There may be inertia among many evaluators to change the ways in which they currently conduct research evaluations for the more challenging, less elegant and more participatory steps just suggested. Those willing to pick up the challenge and experiment with change, however, are likely to enjoy considerable rewards in terms of increased interest in, and appreciation for, the processes engaged in and the findings arrived at.

5.3. Attend to the pathways through which impacts are achieved

As is the case in most international development agencies, those engaged in agricultural research seeks to accomplish their goals via numerous and highly complex pathways. These involve many organizational levels within the system, including international, regional and national research and development organizations, various governmental and non-governmental bodies, agricultural producers, market agents, and consumers. In such a complex setting, evaluations need to provide accounts not only of success, but of how and why both success and failures have occurred. Useful accounts cannot be produced by input–output evaluations that obscure processes within black boxes. From a study designed to assess the impacts of the broadbed maker plough in Ethiopia, using economic surplus methodology, [Rutherford et al. \(2001\)](#) conclude that the entire range of adoption constraints (including farmers' risk-management strategies, the role of extension, changes in household labour demands, farmer modification of the technology, and interdependencies within and between households) all need to be understood if effective means are to be found to help alleviate rural poverty and hunger.

When the findings of impact assessments show that anticipated impacts have not been achieved, it may be unclear where the fault lies—in the policy, the program or in the implementation process ([Mayne, 2001](#); [Berk and Rossi, 1990](#)). As noted by [Pingali \(2001\)](#) and [Bennett \(2000\)](#), the impact assessment methods currently used in the field of international agricultural research are not always able to provide convincing evidence of attribution.

If the international agricultural research centres, and other development agencies, were to undertake more theory-based evaluation, interested stakeholders and potential information users would learn why specific policies and programs, or the activities undertaken to implement them, do or do not work as intended. They would acquire insights into how to improve both the design and the implementation of policies and programs. Finally, they would learn whether, and to what extent, the policy or program was responsible for the outcomes.

5.4. Address organizational and institutional issues

It follows from the discussion of pathways in the previous section that institutions and organizational arrangements play key roles in accomplishing the results sought after by international development agencies. To date, that role is poorly understood, and there has been little methodological and practical work in the area of institutional impact assessment in the context of agricultural

research (IAEG, 2001). Within the CGIAR, it has been noted that “research methods and institutional aspects of research organizations tend to escape any kind of impact analysis” (IAEG, 2001: 23). Maredia et al. (2000: 17), referring to the recent interest in “institutional impacts” within the CGIAR, report that “none of these efforts has been successful in putting economic values on the outcomes of capacity building”. The authors of this paper are not familiar with any studies that have attempted to do this. However, putting credible economic values on outcomes may not be feasible without a better understanding of the ways in which capacity building takes place.

Attempting economic impact assessment at this point may be putting the cart before the horse. There is uncharted ground separating institutional and organizational assessment from economic impact assessment. From the perspective of an economic impact assessment specialist, institutional impact studies are often considered deficient in their concepts, methods and data sources. On the other hand, organizational and institutional specialists are dissatisfied with economic impact assessment methodologies which treat organizations as black boxes or reduce their complexity to “assumptions of stability and equilibrium, of linearity in the relationship between variables, and of proportionality of change in response to causal influences” (Sanderson, 2000: 442).

Multidisciplinary teams, involving specialists in economic impact assessment, institutional analysis and organizational science could explore new and more useful ways of examining and understanding capacity development and the role of organizations in international research and development. Porter (1990), writing about his approach to understanding the competitive advantage of firms, governments and nations, has said that in studying national economic success, there has been a tendency to limit the disciplinary perspective so as to “gravitate to clean, simple explanations and to believe in them as an act of faith” (p.29). He suggests that it would be advantageous to “cut across disciplines and examine more variables in order to understand how complex and evolving systems work” (p.30). By being more open to interdisciplinary impact assessments and other kinds of evaluations, international development agencies would be able to address a broader range of questions of direct interest to a broader range of potential users and thereby increase the practical relevance of their evaluation work to interested stakeholders.

6. Concluding remarks

Impact assessment and other types of evaluation are costly endeavours. It might be argued that adoption of the approaches discussed in this paper to enhance the utility and use of impact assessment and evaluation would be too complex and costly to be practical. However, we also need to consider the costs of *not* engaging in utilization-oriented evaluation as well as the benefits that could accrue from doing technically sound evaluations that serve the needs of identified users.

Acknowledgements

An earlier version of this paper was presented at the international conference “Why Has Impacts Assessment Research Not Made More of a Difference?” held 4–7 February, 2002 in San José, Costa Rica. The conference was co-organized by the Standing Panel on Impact Assessment (SPIA) of the Consultative Group for International Agricultural Research (CGIAR) and by the Economics Program of the International Maize and Wheat Improvement Centre (CIMMYT). The authors thank Ruth Meinzen-Dick, International Food Policy Research Institute, and Stephen Biggs, School of Development Studies, University of East Anglia, for suggestions on an earlier version of this manuscript, and Viviana Galleno MBA, ISNAR, for helping guide us into relevant literature in the extensive field of organizational management.

References

- Balthasar, A., Rieder, S., 2000. Learning from evaluations: effects of evaluation of the Swiss Energy 2000 Programme. *Evaluation* 6 (3), 245–260.
- Bennett, A., 2001. A Member’s Perspective and Expectations with Respect to Impact Assessment in the CGIAR. In: TAC Secretariat (Ed.), *The future of impact assessment in the CIGAR—needs, constraints and options*. Proceedings of a workshop organized by the Standing Panel on Impact Assessment of the Technical Advisory Committee, 3–5 May 2000, FAO, Rome. Technical Advisory Committee of the CIGAR. Food and Agriculture Organization of the United Nations, Rome.
- Berk, R., Rossi, P., 1990. *Thinking about Program Evaluation*. Sage Publications, Newbury Park.
- Bickman, L. (Ed.), 1987. *Using Program Theory in Evaluation*. New Directions for Program Evaluation Series No. 33.
- CGIAR Secretariat, 1998. *Third System Review of the CGIAR*. Secretariat of the Consultative Group on International Agricultural Research (CGIAR), Washington, DC.
- Chelimsky, E., 2001. What evaluation could do to support foundations: a framework with nine component parts. *The American Journal of Evaluation* 22 (1), 13–28.
- Chen, H.T., 1990. *Theory-Driven Evaluation*. Sage Publications, Thousand Oaks, CA.
- Dodgeson, M., 1993. Organizational learning: a review of some literatures. *Organization Studies* 14 (3), 375–394.
- Fetterman, D., 2001. *Foundations of Empowerment Evaluation*. Sage Publications, Thousand Oaks, CA.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, W., Scott, P., Trow, M., 1994. *The New Production of Knowledge*. Sage Publications, London.
- Gibbons, M. Science’s new social contract with society. *Nature* 402 (suppl.) December 2, 1999. C81–C84.
- Goldsmith, A., 1993. Institutional development in national agricultural research: issues for impact assessment. *Public Administration and Development* 13 (3), 195–204.
- Greene, J., Caracelli, V. (Eds.), 1997. *Advances in Mixed-Method Evaluation*. New Directions for Evaluation. No. 74.
- Guba, E.G., Lincoln, Y., 1981. *Effective Evaluation: Improving the Usefulness of Evaluation Results Through Responsive and Naturalistic Approaches*. Jossey Bass, San Francisco, CA.
- Horton, D., 1986. Assessing the impact of international agricultural research and development programs. *World Development* 14 (4), 453–468.
- Horton, D., 1998. Disciplinary roots and branches of evaluation: Some lessons from agricultural research. *Knowledge and Policy* 10 (4), 32–69.
- Horton, D., Mackay, R., 1998. Assessing the organizational impact of development cooperation: a case from agricultural R&D. *The Canadian Journal of Program Evaluation* 13 (2), 1–28.

- Horton, D., Ballantyne, P., Peterson, W., Uribe, B., Gapasin, D., Sheridan, K. (Eds.), 1993. *Monitoring and Evaluating Agricultural Research: A Sourcebook*. CAB International, Wallingford, UK.
- Horton, D., Mackay, R., Andersen, A., Dupliech, L., 2000. *Evaluating Capacity Development in Planning, Monitoring and Evaluation: A Case Study from International Agricultural Research*. ISNAR Research Report No. 17. The Hague: International Service for National Agricultural Research.
- Horton, D., Dupliech, L., 2001. Towards an integrated planning, monitoring, and evaluation system. In: Gijssbers, G., Janssen, W., Hambly, H. (Eds.), *Planning Agricultural Research: A Sourcebook*. International Service for National Agricultural Research (ISNAR), The Hague.
- Hoy, W., Miskel, C., 1996. *Educational Administration: Theory, Research, and Practice*. McGraw-Hill, New York.
- Huber, G.P., 1991. Organizational learning: the contributing processes and literatures. *Organization Science* 2 (1), 88–115.
- Hudson, J., Mayne, J., Thomlison, R., 1992. *Action-Oriented Evaluation in Organizations: Canadian Practices*. Wall & Emerson, Toronto.
- IAEG, 2001. Impact assessment of agricultural research: context and state of the art. In: TAC Secretariat (Ed.), *The Future of Impact Assessment in the CGIAR: Needs, Constraints and Options*. Proceedings of a Workshop Organized by the Standing Panel on Impact Assessment of the Technical Advisory Committee, 3–5 May 2000, FAO, Rome. Technical Advisory Committee of the CGIAR. Food and Agriculture Organization of the United Nations, Rome.
- iSC Secretariat, 2002. Report to the Annual General Meeting of the CGIAR from The Standing Panel on Impact Assessment (SPIA) of the interim Science Council. SDR/iSC: IAR/02/29. interim Science Council of the CGIAR, Rome.
- Joint Committee on Standards for Educational Evaluation, 1994. *The Program Evaluation Standards*. Sage Publications, Thousand Oaks, CA.
- Kumar, K., 1995. Measuring the performance of agricultural and rural development programs. In: Picciotto, R., Rist, R. (Eds.), *Evaluating Country Development Policies and Programs: New Approaches for a New Agenda*. New Directions for Evaluation No. 67. Jossey Bass, San Francisco.
- Lindblom, C.E., 1980. *The Policy-making Process*, 2nd ed. Prentice Hall, Englewood Cliffs.
- Long, B., Kishchuk, N., 1997. Professional certification: a report to the National Council of the Canadian Evaluation Society on the experience of other organizations. Canadian Evaluation Society, Ottawa. Available at: <http://evaluationcanada.ca/txt/longkishchukreport.pdf>.
- Love, A.J., 1991. *Internal Evaluation: Building Organizations from Within*. Sage Publications, Newbury Park, CA.
- Mackay, R., Horton, D., Dupliech, L., Andersen, A., 2002. Evaluating organizational capacity development. *The Canadian Journal of Program Evaluation* 17 (2), 121–150.
- Maredia, M., Byerlee, D., Anderson, J., 2000. Ex Post Evaluations of Economic impacts of Agricultural Research Programs: A Tour of Good Practice. Paper presented to the Workshop on “The Future of Impact Assessment in CGIAR: Needs, Constraints, and Options”, Standing Panel on Impact Assessment (SPIA) of the Technical Advisory Committee, Rome, 3–5 May 2000, Rome.
- Mayne, J., 2001. Addressing attribution through contribution analysis: using performance measures sensibly. *The Canadian Journal of Program Evaluation*. 16(1).
- Mayne, J., 1992. Enhancing evaluation use. In: Hudson, J., Mayne, J., Thomlison, R. (Eds.), *Action-Oriented Evaluation in Organizations: Canadian Practices*. Wall & Emerson, Toronto, pp. 285–287.
- Ndiritu, C.G., 2000. A NARS Perspective on Impact Assessment in the CGIAR. Standing Panel on Impact Assessment (SPIA) of the Technical Advisory Committee, Rome, 3–5 May 2000, Rome.
- Özgediz, S., 1999. Evaluating research institutions: lessons from the CGIAR. *Knowledge, Technology & Policy* 11 (4), 97–113.
- Patton, M.Q., 1997. *Utilization-Focused Evaluation: The New Century Text*. Sage Publications, Thousand Oaks, CA.
- Patton, M.Q., 2002. *Qualitative Research and Evaluation Methods*, third ed. Sage, Thousand Oaks, California.
- Picciotto, R., 1995. Evaluation and development. In: Picciotto, R., Rist, R. (Eds.), *Evaluating Country*

- Development Policies and Programs: New Approaches for a New Agenda. *New Directions for Evaluation* No. 67. pp. 13–26.
- Picciotto, R., 2002. *The Logic of Mainstreaming: A Development Evaluation Perspective*. Operations Evaluation Department. The World Bank, Washington, DC.
- Pingali, P., 2001. Milestones in impact assessment research in the CGIAR, 1970–1999. In: TAC Secretariat (Ed.), *The future of impact assessment in the CIGAR—needs, constraints and options*. Proceedings of a workshop organized by the Standing Panel on Impact Assessment. Technical Advisory Committee, 3–5 May 2000, FAO, Rome. Technical Advisory Committee of the CIGAR. Food and Agriculture Organization of the United Nations, Rome.
- Porter, M., 1990. *The Competitive Advantage of Nations*. The Free Press, New York.
- Rossi, P., 1987. The iron law of evaluation and other metallic rules. *Research in Social Problems and Public Policy* 4, 3–20.
- Russ-Eft, D., Preskill, H., 2001. *Evaluation in Organizations. A Systematic Approach to Enhancing Learning, Performance, and Change*. Perseus Publishing, Cambridge.
- Rutherford, A., Odero, A., Kruska, R., 2001. The role of the broadbed maker plough in Ethiopian farming systems: an ex post impact assessment study. ILRI Impact Assessment Series 7. ILRI (International Livestock Research Institute), Nairobi, Kenya.
- Sanderson, I., 2000. Evaluation in complex policy systems. *Evaluation* 6 (4), 433–454.
- Scriven, M., 1993. Hard-won lessons in program evaluation. *New Directions for Program Evaluation* 58, 1–107.
- Scriven, M., 1998. Empowerment evaluation examined. *Evaluation Practice* 18 (2), 165–175.
- Scriven, M., 1967. The methodology of evaluation. In: Stake, R.E. (Ed.), *Curriculum Evaluation*. American Educational Research Association Monograph Series on Evaluation. Rand McNally, Chicago, pp. 39–83.
- Sonnichesen, R., 2000. *High Impact Internal Evaluation*. Sage Publications, Thousand Oaks, CA.
- TAC Secretariat, 2001. *The future of impact assessment in the CIGAR—needs, constraints and options*. Proceedings of a workshop organized by the Standing Panel on Impact Assessment. Technical Advisory Committee, 3–5 May 2000, FAO, Rome. Technical Advisory Committee of the CIGAR. Food and Agriculture Organization of the United Nations, Rome.
- Thoening, J., 2000. Evaluation as usable knowledge for public management reforms. *Evaluation* 6 (2), 217–229.
- Torres, R., Preskill, H., Piontek, M., 1996. *Evaluation Strategies for Communicating and Reporting: Enhancing Learning in Organizations*. Sage Publications, Thousand Oaks, CA.
- Weiss, C., 1999. The interface between evaluation and public policy. *Evaluation* 5 (4), 468–486.
- Weiss, C., 1998. *Evaluation*. Prentice Hall, Upper Saddle River, NJ.
- Whitmore, E. (Ed.), 1998. *Understanding and Practicing Participatory Evaluation*. *New Directions for Evaluation* No. 80. Jossey Bass, San Francisco.
- Winberg, A., 1992. Effective communication (chapter 17). In: Hudson, J., Mayne, J., Thomlison, R. (Eds.), *Action-Oriented Evaluations in Organizations: Canadian Practices*. Wall and Emerson, Toronto.
- Worthen, B., Sanders, J., Fitzpatrick, J., 1997. *Program Evaluation: Alternative Approaches and Practical Guidelines*. Longman, New York.