The Canadian Journal of Program Evaluation Vol. 17 No. 2 Pages 121–150
ISSN 0834-1516

EVALUATING ORGANIZATIONAL CAPACITY DEVELOPMENT

Ronald Mackay
Concordia University
Montreal, Quebec

Douglas Horton
International Service for National Agricultural Research
The Hague, The Netherlands

Luis Dupleich
Government of Bolivia
La Paz, Bolivia

Anders Andersen
Danish Ministry of Foreign Affairs
Copenhagen, Denmark

Abstract: While substantial sums are being invested in the development of organizational and institutional capacities, the design and management of capacity development efforts leave much to be desired. Few capacity development initiatives have been systematically and thoroughly evaluated. This article describes the conceptual frameworks and methods used to evaluate a multi-site, regional capacity-development project in Latin America and the Caribbean undertaken to strengthen planning, monitoring, and evaluation in agricultural research organizations. The article discusses some of the challenges facing capacity development and its evaluation, outlines the procedures employed, and illustrates these with some consolidated findings in response to four evaluation questions: What were the main contributions of the project to agricultural research management? How were the results achieved? What factors facilitated their achievement? and What lessons can we learn to improve future capacity development efforts and their evaluation?

Résumé: Bien que des sommes substantielles soient investies dans la mise en valeur du potentiel institutionnel et organisationnel, la conception et la gestion des efforts de renforcement des capacités laissent fortement à désirer. Par ailleurs, rares sont les efforts de mise en valeur du potentiel qui ont fait l'objet d'une évaluation systématique et approfondie. Cet article décrit les cadres conceptuels et les méthodes utilisés par les évaluateurs d'un

projet régional de renforcement des capacités mis en œuvre à plusieurs sites en Amérique latine et aux Caraïbes. L'article examine certains des défis en matière de renforcement des capacités et son évaluation, décrit les grandes lignes des procédures adoptées, les illustre avec quelques exemples et présente les résultats en réponse aux quatre questions suivantes: Quelles sont les principales contributions du projet à la gestion de la recherche? Comment les résultats ont-ils été atteints? Quels facteurs ont facilité leur mise au point? et Quelles leçons peut-on tirer de l'expérience?

FOCUS OF THE EVALUATION

This article reports on the evaluation of a six-year, multi-site, capacity-development project designed to strengthen the planning, monitoring, and evaluation (PM&E) capabilities of national agricultural research organizations (NAROs) in Latin America and the Caribbean. The project was initiated by the International Service for National Agricultural Research (ISNAR), one of 16 independent institutes affiliated with the Consultative Group on International Agricultural Research (CGIAR).

Policy makers, managers, and those responsible for development cooperation continually search for improved ways of strengthening organizational capacity. Capacity refers to both the organizational arrangements and the technical capabilities that permit organizations to carry out their primary functions and thereby accomplish their development goals. Of particular interest to the evaluation community is the need for methods to evaluate capacity development efforts (Horton, 2001; Jackson & Kassam, 1998; Moore, 1995) and to use evaluation to foster organizational development (Rowe, 1999).

This evaluation was undertaken to assess the contributions made by a capacity development project to critical management functions in participating organizations, how these were achieved, and what lessons might be drawn from the exercise. This article briefly reviews organizational capacity development as it is currently viewed by development organizations. It then describes the capacity development effort itself. Next, it explains the policy context of the development effort and describes the conceptual frameworks and methods employed in the evaluation. The results of the evaluation are presented briefly to illustrate aspects of the evaluation methodology. Some lessons are also reported.

ORGANIZATIONAL CAPACITY BUILDING AS A DEVELOPMENT THRUST

Building the capacity of organizations in the developing world is a goal shared by many development agencies. However, aid agencies (including the World Bank and USAID), partners, and beneficiaries are finding organizational capacity development more elusive than merely delivering or acquiring resources (Austin, 1993; Moore, 1995).

Defining "Capacity Development"

Analysts tend to define capacity-related problems using the concepts and terms of their own disciplines. In the field of agricultural research, economists tend to see capacity requirements in terms of policy research and the remedies to lie in regulatory and fiscal mechanisms. Biological scientists see capacity problems in terms of gaps in scientific expertise or technical resources and the solution to lie in opportunities for advanced study and the upgrading of technical facilities. Specialists in the organizational sciences tend to view capacity problems more comprehensively in terms of the systems and subsystems that make up the organization and focus attention on organizational culture, and management practices and processes (Amit & Schoemaker, 1993; Grant, 1995; Prahalad & Hamel, 1990).

In this evaluation, we use the term *organizational capacity development* to refer to the processes of organizational learning and change in planning, monitoring, and evaluation engaged in by the PM&E project team and the R&D organizations, at both the individual and higher organizational levels, intended to help maintain their viability and improve their operations.

The Role of Evaluation in Capacity Development

The ultimate impact of capacity-development programs depends upon the appropriate use of evaluation. Those who design programs need to review the existing capabilities and identify important areas that require strengthening. Managers need to monitor activities and evaluate results in order to adjust, redirect, and improve the effectiveness of their organizations' efforts. They also need to learn, from post-hoc evaluations, about the strengths and weaknesses of these efforts. Finally those who fund capacity-development initiatives need information about their results and impacts in order to justify their continued support.

Few capacity-development initiatives have systems for monitoring or evaluating changes in the organizations they are designed to strengthen, and there are few recognized methods to evaluate their processes, outputs, and impacts (Rist, 1995; Taschereau, 1998). It is difficult and costly to evaluate the contributions of any type of program. In economics and the social sciences, impact assessment is viewed as requiring a rigorous experimental or quasi-experimental research design that allows precise measurement of treatment effects (Alston, Norton, & Pardey, 1995; Rossi, Freeman, & Lipsey, 1999). Such designs have been successfully applied in the measurement of impacts of agricultural research, public health, and some other types of public programs. However, their application to organizational capacity-development programs is highly problematic.

World Bank evaluators have noted that, in the sphere of agricultural research and development, some types of impact assessment are more difficult than others, because of the varying complexity and confounding factors involved (Anderson & Dalrymple, 1999, pp. 41–42). Assessing the impact of new crop varieties, in terms of increase in yield per acre, is perhaps the least difficult. Assessing impacts of research on crop management is more complex. Dealing with the contributions of participatory research approaches adds additional complexity. And one of the most complex of all, they claim, is the assessment of capacity development in research organizations.

Researchers have listed many problems facing the evaluation of organizational capacities and their development (Maconick & Morgan, 1999; Moore, 1995). These include: (a) capacity development itself has a diffuse and often poorly defined concept; (b) organizational capacity development is generally considered not to be a goal in itself, but a means to other development goals; (c) capacity development processes are difficult to specify and isolate and have few in-built mechanisms to draw attention to poor progress; (d) the attribution problem is especially acute in the case of capacity development where results may emerge only over a long period of time; (e) organization and management studies, the disciplines most relevant to capacity development, are only beginning to create a theoretically well-founded and commonly accepted body of concepts and terms; and (f) despite the existence of several frameworks for assessing organizations and capacity-development programs (Lusthaus, Anderson, & Murphy, 1995; Montague, 1997; Taschereau, 1998), there are few reports of their practical application.

The Policy Context of the Evaluation

A major obstacle facing the promise of agricultural research to contribute to human welfare, food security, and sustainable environmental management is the capacity to manage research and development organizations effectively (Horton, 1986). The need to strengthen management capacity has been noted in many types of organizations (Hilderbrand & Grindle, 1995; Hudson, 1999). An organization's performance depends not only upon its financial, human, and other resources but also on its managers' capacity to deploy these resources in the successful pursuit of its strategic goals. Within the international agricultural research community, ISNAR has the mandate to strengthen the management of national agricultural research organizations in the developing world. It accomplishes this task by conducting applied management research, disseminating information, providing management training, and collaborating with its clients in diverse aspects of organizational change. Management tools and approaches have been developed by ISNAR for use in such areas as strategic management, (Gálvez, Novoa, de Souza Silva, & Villegas, 1995); planning (Gijsbers, Janssen, Hambly Odame, & Meijerink, 2001), program monitoring and evaluation (Bojanic, Hareau, Posada, Ruíz, & Solís, 1995; Granger, Grierson, Quirino, & Romano, 1995), and organizational performance assessment (Peterson & Perrault, 1998).

OBJECT OF THE EVALUATION

ISNAR's PM&E capacity development project was implemented from 1992 to 1997. The aim was to establish integrated PM&E systems and thereby enhance research management in participating organizations. Improvements in management were expected to contribute to the relevance, effectiveness, and efficiency of agricultural research programs, to the production of new information and technology, and to its use in the productive sector. In order to achieve its goals, the project carried out activities in three areas:

- *Information.* Reference books and training materials were prepared for use in training events and workshops and for distribution to managers and libraries throughout Latin America. Prepared in Spanish, the materials were later translated into English and distributed in other regions.

- *Training and workshops.* A regional group of trainers was established, and its members organized and delivered a

number of sub-regional training events. Several regional workshops were also organized to plan and review the project's activities and to disseminate its results to high-level managers in the region.

- *Facilitation of organizational change.* From 1992 to 1995, the project focused on information and training activities. In 1996, it began providing direct support for organizational change processes in selected organizations that were committed to making significant improvements in their PM&E systems. These "pilot cases" were in Costa Rica, Cuba, Panama, and Venezuela.

The project's strategies and activities were guided by three basic values. *Active participation* of the project's intended beneficiaries was considered essential for ensuring the relevance of its activities, commitment to its objectives, and application of its results. *Learning by doing* was viewed as an essential process of organizational strengthening. To foster learning, staff members of ISNAR and participating organizations jointly planned activities, tested innovations, reviewed results, and modified plans accordingly during project implementation. *Respect for diversity* ensured that the project staff tried to understand and build upon the wide range of knowledge and experience present in the region and sought to work with local managers to develop locally adapted solutions to management problems.

PLANNING FOR THE EVALUATION

The evaluation effort was participatory from start to finish. At the outset, meetings were organized with representatives of major stakeholder groups engaged in the capacity development effort and potential users of the evaluation results. These served to determine (a) the purposes and scope of the evaluation, (b) the key questions upon which it would focus, (c) the kinds and sources of data required to answer these questions, and (d) the most suitable data collection strategies and instruments. An advisory committee was established, made up of four individuals with broad international evaluation expertise, to provide feedback and guidance at critical points throughout the evaluation. Later meetings allowed participants to review data, results, and findings, to share draft reports, and to synthesize feedback. A final meeting allowed stakeholders, including the donors and specialists in capacity development, to assemble information for both the final report (Horton, Mackay, Andersen, & Dupliech, 2000) and a meta-evaluation (Murphy, 2000).

Focus via Evaluation Questions

The evaluation was directed by four key questions that addressed the results of the capacity development effort, the processes and strategies employed by both national organizations and ISNAR, the factors facilitating successful outcomes, and the lessons learned about the execution and evaluation of capacity development efforts.

Concepts and Frameworks Used to Guide the Evaluation

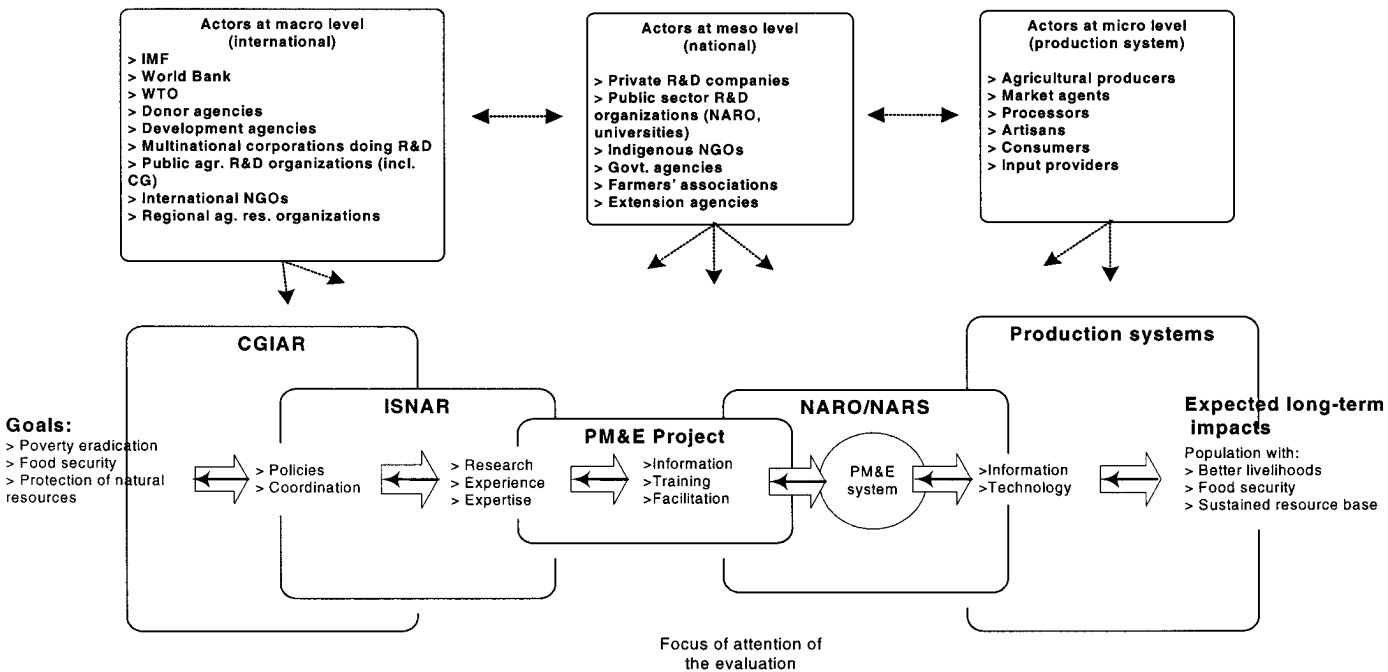
Various conceptual frameworks were employed to structure the evaluation.

Impact chain. A hypothesized "impact chain" (Figure 1) originates with the goals of the CGIAR and ends with the anticipated long-term impacts on poverty, food security, and the environment. The chain includes links corresponding to the CGIAR system, ISNAR, the PM&E project, a national agricultural research organization, and an agricultural production system. The scope and focus of the present evaluation is identified within the chain. It is important to stress that the evaluation reported here does not address the entire impact chain. It restricts itself to the relatively little-explored link between a capacity development effort and the organizations it works to strengthen. Many different actors and factors affect agricultural innovation processes and the resulting social, economic and environmental effects (Biggs, 1990; Röling & Wagenmakers, 1998). Mayne (2001) has made the same observations with regard to the impacts of government programs. Attention has been drawn to the difficulties inherent in measuring the impacts of a single project on socioeconomic variables at the level of the ultimate intended beneficiaries (Anderson & Dalrymple, 1999).

The Project Logic Model. Capacity-building projects are goal-directed. Activities, processes, and interventions are undertaken in the belief that they will enhance organizational capacities in the areas where needs have been diagnosed. The assumptions or hypotheses that link project activities and desired outcomes represent the logic directing program action. The hypotheses and assumed causal relationships between program components may be more or less explicit and more or less consciously informed by theories about organizational behaviour.

A simple logic model for the PM&E project captures the assumptions that link the project's activities, outputs, purpose, and ulti-

Figure 1
"Impact chain" from CGIAR development goals to expected long-term impacts, locating the focus of attention of the evaluation

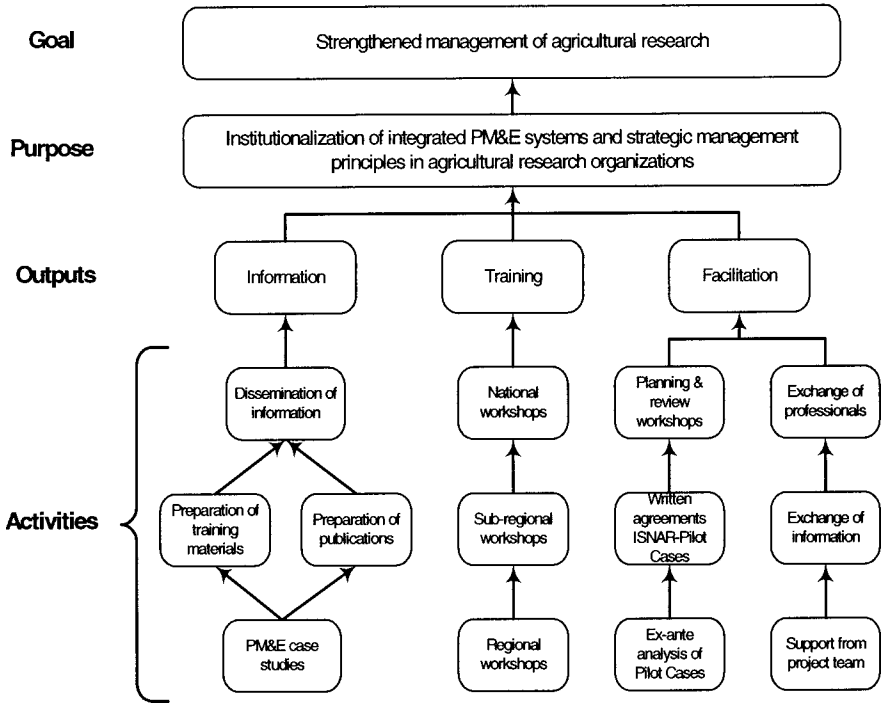


Note. The rectangles representing organizational units and systems overlap to reflect the co-production of key outputs. The project provides a crucial link betwen ISNAR and national agricultural research organizations and systems.

mate goal (Figure 2). As noted earlier, in order to strengthen agricultural research management the project employed three types of activities, grouped in the "project components" of information, training, and facilitation of change processes.

Organizational assessment framework. Frameworks for diagnosing or assessing organizations generally present a model of the organization that represents it as an open system and provides checklists to guide the collection and analysis of information related to key variables. The evaluation reported here employs an organizational assessment framework developed by Universalia and IDRC in Canada (Lusthaus et al., 1995). It was selected for this evaluation because of its compatibility with current organizational theory, comprehensiveness, flexibility, and relative simplicity. It had proved its value in an earlier evaluation conducted by members of the ISNAR evaluation team (Horton & Mackay, 1998).

Figure 2
Hierarchy of PM&E project objectives



This framework views an organization's performance as a function of its operational environment, its organizational motivation, and its organizational capacity. *Operational environment* refers to the legal, social, and economic context in which the organization operates. *Organizational motivation* refers to internal factors that influence the direction of the organization and the energy displayed in their activities; these are influenced by organizational culture, incentives, and similar variables. *Organizational capacity* refers to the staff complement and resources possessed by the organization as well as its structure, management processes, and systems, and linkages with other organizations. Organizational performance is gauged in terms of the organization's effectiveness, efficiency, relevance, and sustainability. *Effectiveness* refers to the degree to which the organization achieves its goals; *efficiency* refers to the degree to which unit costs are minimized; *relevance* refers to the extent to which the organization's outputs and results are valued by its stakeholders; and *sustainability* is achieved through effective acquisition and development of financial, human, and physical resources.

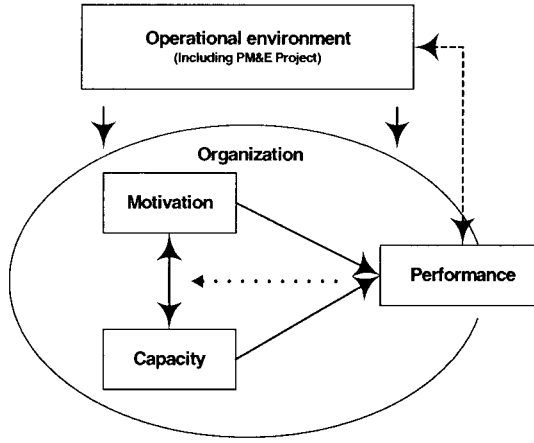
Because the PM&E project sought to bring about changes by working with agricultural research managers, the framework was adapted and applied at two levels: the individual participant and the organization as a whole.

In this expanded framework, a capacity-development initiative is viewed as one contributing element in the operating environment of the individual or the organization. It can have *direct* effects on other factors in the environment and on the individual's or the organization's motivation and capacity. Through its effects on the environment, motivation, or capacity, a development intervention can *indirectly* contribute to the performance of the individual or the organization. The relationship among the four dimensions and examples of critical factors associated with each dimension are illustrated in Figure 3.

Integrated Evaluation Framework

Based on the PM&E project's theory of action and the assessment framework just described, an integrated evaluation framework was developed that relates each of the project components to four potential areas of impact at the level of individuals and organizations (Figure 4). This framework provides a visual representation of a set of complex relationships that can otherwise be elusive.

Figure 3
Organizational assessment framework



Operational environment. The external environment in which the organization carries out its activities. Examples:
- Administrative and legal systems in which the organization operates
- Political environment
- Technological options
- Social and cultural environment
- Economic trends
- Stakeholders

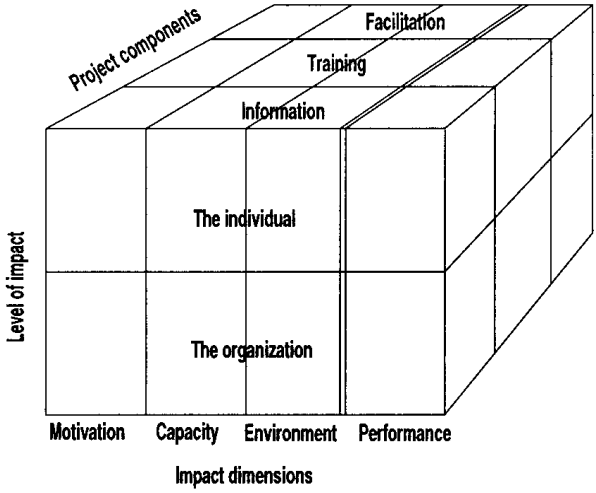
Motivation. Refers to internal factors that influence the direction of the organization and the energy displayed in its activities. Examples:
- The organizational culture
- Incentive and rewards systems
- The institutional "climate" in general
- The history and traditions of the organization
- Leadership and management style
- A generally recognized and accepted mission statement
- Performance-related incentive plans
- Shared norms and values promoting teamwork toward organizational goals

Capacity. The resources, knowledge, and processes employed by the organization. Examples:
- Strategic leadership
- Organizational structure
- Human resources
- Financial resources
- Physical infrastructure
- Program process management
- Inter-institutional linkages

Performance. Is gauged in terms of four key indicators:
- Effectiveness: The degree to which the organization achieves its objectives
- Efficiency: The degree to which it generates its products using a minimum of inputs
- Relevance: The degree to which the organizational objectives and activities reflect the necessities and priorities of key stakeholders
- Financial sustainability: The conditions to make an organization financially viable include multiple sources of funding, positive cash flow, and financial surplus

Source: Developed from Lusthaus, Anderson, and Murphy (1995).

Figure 4
Integrated evaluation framework



Note. The three project components are assumed to contribute mainly to aspects of individual and organizational motivation, capacity, and the environment. The double line between environment and performance reflects that performance is a function of motivation, capacity, and environmental variables. Hence, the project's contributions to performance are indirect.

FIVE COMPLEMENTARY EVALUATION STUDIES

Five evaluation studies were designed to assess the impact of the three project components on the environment, motivation, capacity, and performance of participating individuals and organizations. The studies involved the use of diverse sources of information reflecting the multiplicity of stakeholders who might be affected by the project and its results. They also involved multiple methods for data collection and subsequent analysis (Table1).

Study findings were triangulated using several different sources of data (e.g., organizational records, observations, and different classes of informants), several different researchers, and different methods (mail-out questionnaires with closed- and open-ended questions, personal and group interviews, and review of documents). Specific steps were also taken to enhance the validity of results within specific studies.

Study 1. The ISNAR PM&E Capacity Development Project. This study, undertaken by the PM&E project team assisted by a member of the evaluation team (Cheaz, de Souza Silva, Andersen, & Horton, 1999), provided a descriptive review of the project since its inception. It presents background information on the institutional setting of the project and outlines its objectives, strategies, activities, and products. The sources and uses of project resources are also documented. The report was drawn from PM&E project records. A member of the evaluation team searched the records of the PM&E project for the period under scrutiny and referenced the sources from which all results were drawn. Two members of the project team audited the results and the sources used.

Studies 2 and 3. Impacts of information and training. There have been relatively few evaluations of the impact of information on organizational capacity or performance (CTA, 1998). The information study was carried out to evaluate the distribution and use of the

Table 1
Evaluation matrix: The five complementary evaluation studies

Study	Objectives	Methods	Sources of data
Study 1: The ISNAR PM&E capacity development project	Review the project's objectives, strategies, activities, and outputs	Self-assessment	Project records
Study 2: Impacts of information	Analyze dissemination, use, and impact of publications	Mail survey	500 recipients of project publications from 140 organizations in 24 countries
Study 3: Impacts of training	Analyze impacts of training	Mail survey	150 training participants from 60 organizations in 20 countries
Study 4: Changes in PM&E in the pilot cases	Analyze changes in PM&E in the pilot cases; identify contributions of the PM&E project; determine effects of the changes on organizational performance	Facilitated self-assessment	Collaborators in 3 pilot cases
Study 5: Dynamics of PM&E in Latin America and the Caribbean	Analyze changes in PM&E in the region; identify contributions of the PM&E project; determine effects of the changes on organizational performance	Case studies	Informants, documents, and observations in 9 organizations in 8 countries

project's publications and the effects on individuals and their organizations. A postal survey was employed to collect information from all known individuals who had received project publications. Respondents provided information on the use of project publications and on their usefulness relative to other publications on the same topics. They also scored the impact of project publications on a set of indicators that was used to determine the degree to which information had affected the motivation, capacity, operational environment, and performance of these professionals and their organizations.

There is an extensive literature on the evaluation of training. A widely used model identifies possible effects of training at four levels: participants' reaction to the training; participants' learning as a result of the training; change in participants' behaviour resulting from the training; and the subsequent impact on the organization as a result of participants' behaviour change (Kirkpatrick, 1998). Study 3 was concerned primarily with the impact of training on participants' behaviour and on their organizations. A questionnaire was mailed to all professionals in the region who had participated in project training activities. As in the information study, respondents scored the impact of training. A standard set of indicators was developed to represent the degree to which the training had affected motivation, capacity, operational environment, and performance of these professionals and their organizations. In addition, open-ended questions solicited concrete examples of the results of training at both individual and organizational levels.

Mailing records were used to identify 500 recipients of project publications from 140 organizations in 24 countries and 150 recipients of project training in 60 organizations in 20 countries in Latin America and the Caribbean. In both studies, each individual was sent a mail survey. Two months later, the addresses of non-respondents were updated where possible and a second copy of the survey with a reminder letter was sent to them. Where fax numbers were available, the survey was faxed rather than mailed. The final response rate for the publications survey was 29% and for the training survey, 45%.

In a review of methodology textbooks, Goyder (1985) found that expected postal survey response rates ranged from 30% to 70%. The greatest concern with postal surveys is non-response error or selection bias (Bickman & Rog, 1998). To determine if there was a difference between respondents and non-respondents, and how this might affect the validity of the findings, the evaluation team carried out

two analyses exclusively with the training survey. The first was a desk study that compared the profiles of respondents and non-respondents. The second was a telephone survey of non-respondents. The profile generated by the desk study indicated that respondents generally had had more intense, more frequent, and more recent interactions with the PM&E project than non-respondents. However, respondents represented a heterogeneous group of professionals in 27 of the 36 agricultural research organizations that had participated in training over the course of the project.

A random sample of 20 of the 79 professionals who did not respond to the training questionnaire was selected for interview by telephone. Of these, 9 had moved on to different organizations and could not be contacted. Of the 11 who were interviewed, 7 said that they had never received the questionnaire, that it had been misplaced, or that they had received and completed it and returned it in the mail. All 7 spontaneously mentioned the impacts of the PM&E project on them or on their organization. None claimed no impacts to report. Of the remaining three, one claimed the questionnaire did not adequately capture the PM&E situation in his organization, one asserted that because his organization had been the focus of a case study, he did not think it necessary to respond, and one claimed that there were no impacts to report in his organization at the time the questionnaire was received.

Conclusions from these two studies suggest that some overall results (particularly numerical and averaged) are biased toward the "high" end given the fact that respondents are not fully representative of the universe. However, responses came from a wide array of professionals and organizations. The principal reason for not responding appears to be that the questionnaire was not received. The authors believe, therefore, that while the averaged, numerical results stemming form the information and training studies might overestimate the impacts achieved by the PM&E project, it should be possible to draw general conclusions from the training study, especially where the results are corroborated by data from other sources.

A "coupled credibility" technique was used in the publications and training study. Both surveys had asked respondents if PM&E practices in their organizations had been changed, based on information contained in the project's publications (Study 2) and based on training (Study 3). In both surveys, in order to give greater credibility to affirmative answers, respondents were asked to provide concrete examples of the changes they claimed to have resulted. Only affirma-

tive responses supported by concrete examples were counted as instances where project publications or training had resulted in organizational change.

Study 4. Changes in PM&E in the pilot cases. The PM&E project's facilitation component was evaluated by means of self-assessment exercises in the pilot case organizations in which management and staff in each organization analyzed the changes that had taken place in the organization. The self-assessment procedures and instruments were developed by the evaluation team based on a "hermeneutic dialectic process" described by Guba and Lincoln (1989, pp. 72–73, 149–155). The process employed a constructivist methodology permitting those involved in and affected by change in the organizations to explore and develop their understanding of the changes, how they were brought about, and the perceived causal links with the activities of the PM&E project. Findings from the three self-assessments were subsequently synthesized at a workshop involving the change agents from the three pilot cases, members of the PM&E project team, and members of the evaluation team.

Study 5. Dynamics of PM&E in Latin America and the Caribbean. Nine case studies were carried out to document changes in PM&E systems in the region's agricultural research organizations since 1992, in order to determine the contributions of the project to change in PM&E and to identify the effects of changes in PM&E on organizational performance. Results were compared with those of similar case studies conducted in the same organizations five years earlier (Novoa & Horton, 1994). Perceptual data were collected through structured interviews with each organization's managers and staff and with key external stakeholders. These perceptual data were supplemented by a review of each organization's documents (for example, strategic and operational plans for each organization; forms for research project preparation, review, and reporting; progress reports; and final evaluation reports). PM&E facilities and practices were observed directly, to the extent possible during country visits. Each country visit lasted from 5 to 14 days. Data collection was carried out jointly by one or more of the evaluation team members accompanied by one or more members of the organization under study. The case-study investigators prepared separate reports for each case study. A synthesis report was then prepared (Novoa & Horton, 1999).

Despite the frequent use of the term, the "case study" is one of the least understood methods employed by evaluators. In the evaluation reported on here, five different individuals led nine case study teams that involved not only members of the evaluation team but

also one or two members of the organization under scrutiny. For this reason, it was essential that all of those involved understood the purpose of and conducted the studies in the same way. Following Yin (1994) a case-study protocol was developed that went far beyond directions for the common use of the structured interview instrument. The protocol specified the procedures to be employed at every stage from fieldwork to analysis. It included a statement of the overall purpose of the case studies; directions for entry into the organization; the relationship to be established with one or more designated evaluation partners within the organization; the nature and range of the documents to be reviewed; the sample of centres, managers, and researchers to be drawn; how data from interviews, observations, and document review should be recorded and analyzed; and it stipulated that an oral presentation of the principal findings should be made to the senior management team and whomsoever it wished to invite, prior to exit from the organization.

Use of member checks and expert opinion. The "exit meeting" just described served as a "member check" whereby findings and evidence were tested with key members of the organizations (Guba & Lincoln, 1989). At various points during the evaluation process, impacts reported by workshop participants and others familiar with the project were also followed up through telephone conversations, field visits, and correspondence in order to obtain detailed information and check the validity of claims. Two meetings representing the full range of stakeholders were held, in Panama and Ecuador, to review data, results, and findings, to share draft reports, and to obtain feedback. The advisory committee mentioned earlier provided feedback and guidance at critical points during the evaluation. Agricultural research leaders with a sound overview of the region were also asked for their views of the contributions of the project to agricultural research in the region.

FINDINGS

The first task of the evaluation was to ascertain that the project had involved the people in the participating organizations that were most likely to benefit from, promulgate, and help institutionalize PM&E systems. The project was intended to involve research scientists who also carried mid-level management responsibilities within their organizations, in particular, responsibilities associated with the activities of developing program and project plans, monitoring their progress, and evaluating their results. The project logic predicted that this type of participant would be in a favorable position

to replicate the training within their organizations and thereby develop the critical mass with the necessary capabilities to bring about changes in PM&E practices within their organizations.

The profile of those who participated in project training and responded to the questionnaire (45% return rate) shows that on average they spend about 75% of their total work time equally divided between the tasks of management, PM&E, and research. The remaining 25% is taken up with other duties such as extension work. From this information, it can be concluded that the project materials and project training were reaching the targeted participants.

Figure 5
Examples of training impacts at the level of the organization

Respondents reported a total of 212 examples of impacts at the level of the organization. There were 56 examples provided in organizational motivation, 128 in organizational capacity, 12 in the external environment, and 16 in organizational performance. Below is a small but representative selection of examples reported in each of the 4 organizational dimensions.

Impacts on organizational motivation

- Evaluation is viewed more favourably (than previously). It is now seen as part of a substantial organizational learning process. *Chile*.
- A new culture is evolving, one in which the importance and necessity is appreciated of ensuring that the organization has an integrated PM&E system. *Colombia.*
- (Training has promoted) the development of a strategic planning culture. *Dominican Republic.*

Impacts on organizational capacity

- Control activities such as monitoring and evaluation are now conducted in a more integrated and organized way in both research and extension projects. *Argentina.*
- The enhanced development and implementation of strategic planning process at the organizational level and also at the level of the research stations. *Ecuador*.
- Within our region, small working groups have been set up to strengthen understanding and to set the process of establishing PM&E as essential activities. *Saint Vincent and the Granadines.*

Impacts on the environmental dimension of the organization

- From my perspective, an important impact has been the improvement of relations between the various organizations within MINAG and with other organizations within different ministries. All of these organizations have been strengthened as a result of the project workshops and the use of PM&E management tools employed within the strategic approach. *Cuba*.
- In this centre research project planning has been encouraged, resulting in increased financial resources and enhanced links with the private sector. *Argentina*.

Impacts on organizational performance

- An improved response to external demands, especially those based on collaborative undertakings with the private sector. *Mexico*.
- Better regional planning based on confirmed user needs and integrating research, testing, and development. *Argentina.*
- Increased viability of the organization, the extent to which the research agenda is related to user needs, and the capability of the organization to respond to external demands. *Cuba*.

What Have Been the Principal Impacts of the Project?

Impacts of ISNAR's PM&E project were expected within participating organizations and their operating environments (Figure5).

Most of the reported examples of impacts on organizational motivation refer to concrete efforts directed to improve PM&E or to change in the organizational culture toward greater emphasis on performance as demonstrated by the relevance of products.

Most of the reported impacts on organizational capacity fall into four groups:

- Improvements in PM&E procedures
- Development of strategic plans
- Expansion of professional capabilities in PM&E
- Project development in line with the objectives identified in the strategic plan.

Most examples of the project's reported impact on organizational performance refer to the increased responsiveness of research plans and research outputs to the demands of external stakeholders. Impacts on the environment referred mainly to increased interaction and communication with other organizations and to an ethos of increased credibility enjoyed by the organization as a result of its products being recognized as increasingly relevant to producers and other stakeholders.

How Were These Impacts Achieved?

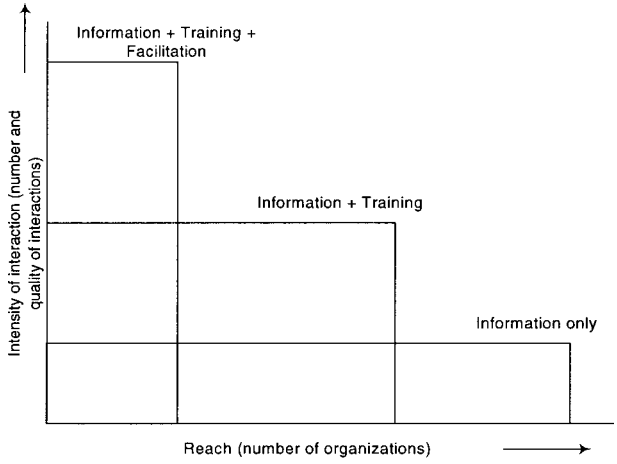
Extrapolating from Montague (1997), capacity-development interventions can be characterized by their reach and the intensity of their interactions with clients or intended beneficiaries. *Reach* refers to the breadth of influence over which an intervention spreads its resources — the number and type of clients it serves. *Intensity of interaction* refers to the amount of time, money, and energy exerted in interacting with clients. The relationship between reach and intensity of interaction of the three project components is illustrated in Figure6.

The project's *information component* had the greatest reach but the least intensity of interaction with the clients. The intention behind the dissemination of the project publications was to offer relevant information about PM&E to the agricultural research and development community at large and, at the same time, to support the training initiatives and the change processes in the pilot cases. Project

publications were received by more than 500 professionals working in some 140 agricultural research organizations around the world. They were distributed in a number of ways: through ISNAR's mailing lists, to individuals who requested them, and to those who participated in project workshops.

The project's training, which entailed directed interaction between the project team and agricultural research managers, reached fewer individuals and organizations in the region. The project provided training in PM&E and management-related topics for approximately 150 middle managers from 60 organizations in Latin America and the Caribbean. Some managers participated in several workshops, first as trainees and later as trainers. The planning and review workshops, which brought together senior managers of organizations in the region, also represented an intermediate level of reach and intensity. Participants learned how PM&E could be a valuable tool for them to manage their organizations. They also shared experiences among themselves and with the project team. In this way, the project benefited from the knowledge and experience of senior managers in the region.

Figure 6
Relationship between reach and intensity of interaction in the three project components



Source: Adapted from Montague (1997).

Pilot case facilitation was a long-term, intensive effort. Here, the project's reach —in terms of the number of organizations — was smaller but the intensity of its interaction with these organizations was far greater. The project provided more copies of the publications and training materials and more opportunities for participation in regional, sub-regional, and national workshops. It also provided direct support for organizational change. Project personnel and external facilitators joined forces with managers and collaborators in pilot case organizations to facilitate strategic planning, to design and implement integrated PM&E systems, and to guide organizational change processes. The intensity of the pilot case interventions is reflected in the large number of people trained in each of these cases.

What Factors Contributed to Impacts?

Environmental factors. ISNAR's PM&E project was conceived and executed at a propitious time. The environment of national agricultural R&D organizations was so turbulent (competition from the private sector, restructuring within the public sector, increased international competition for markets, etc.) and was experiencing such adverse consequences (reduction in core budget allocations from government, the loss of superior personnel to the private sector, etc.) that NAROs were already highly motivated to alter their management practices. Many were actively contemplating change and so were immediately able to use the ISNAR project as a springboard.

Factors within the control of the project. The PM&E project incorporated into its design many factors intended to contribute to its success. These included:

- the project philosophy
- the initial diagnosis of the state of PM&E in the region
- the development of a project logic model
- the training-of-trainers
- the quality of the reference and training materials developed.

The project philosophy. This stressed the active participation of the organizations' researchers and managers throughout the project, ranging from determining goals and focus to taking control of the change process. It also stressed the creation of knowledge using networking techniques to draw upon existing experience and expertise from many disciplines in the region and externally. These princi-

ples built a strong sense of ownership and common commitment to achieving desired results.

The initial diagnosis of the state of PM&E in the region. The diagnosis served to confirm the weaknesses that managers already suspected within their organizations. The opportunity it afforded them to compare their practices with those in similar organizations in Canada and the USA had a galvanizing effect on their desire to build capacity in PM&E. Evidence of weakness in PM&E also helped to ensure that the appropriate middle managers were nominated to participate in the project — those who were in positions to improve and integrate PM&E activities most directly and have them work throughout the organization.

The development of a project logic model. The project logic model was prepared in the early stages of the PM&E project and required a large number of persons involved in the capacity development initiative to discuss and agree upon the results that they wanted to pursue and the means they believed most likely to achieve them. This shared means-ends framework as an instrument representing their project helped to coordinate and align the efforts of participants. It also permitted ongoing adjustments to the plan to be discussed in a principled and rational manner so that everybody could understand why changes were being made and what they meant in terms of their own procedures and work routines.

The training-of-trainers. Using current principles and methods of adult education, training-of-trainers gave impetus to the "snowball" effect that resulted in large numbers of training sessions being delivered within the participating organizations. The participatory development of training materials using regional expertise ensured the relevance and face validity of the instructional content. The intense, residential character of the "training of trainers" workshops ensured devotion to the task without external interruptions. All participants who successfully completed training were awarded certificates of proficiency. The region now counts on a large cadre of trainers in PM&E, some of whom have reported extending their training into the broader university and corporate community.

The quality of the reference and training materials developed. By meeting high standards, the training materials developed for distribution and training contributed to the credibility of the project and therefore the willingness of the organizations to embrace its precepts and information. A monitoring and evaluation "sourcebook" was published jointly by ISNAR and reputable international tech-

nical publishers, in both English and Spanish (Horton, Ballantyne, Peterson, Uribe, Gapasin, & Sheridan, 1993, 1994). The training manuals and materials were developed by experts within the participating organizations under the supervision of an expert in andragogy. These were also subjected to rigorous analysis by external subject-matter experts and experts in training and adult education. The care taken to ensure currency and quality won credibility for project publications not only with those organizations participating in the project, but also with other research organizations in the region and elsewhere in the world.

LESSONS

On reflection and in preparation for this paper, the authors extracted lessons not only about the evaluation of capacity development efforts, but also about their initiation. We start with the latter lessons first.

Lessons Associated with Understanding Capacity Development Initiatives

Start out with an adequate overview of the capacity development initiative within the broader, expanded chain of desired results (see Figure 1 above). It is helpful to understand organizational capacity building efforts on two levels: the policy level and the level of the initiative itself (Berk & Rossi, 1990; Knox, 1996). The first refers to the policy space within which the organization exists and within which the particular effort is being undertaken. For example, in the case of national agricultural organizations, the policy space may include: (a) the achievement of national food security, (b) the reduction of rural poverty, (c) the creation of food surpluses to earn export dollars, or (d) the sustainable management of the physical environment. The program level refers to the organized set of procedures (i.e., the capacity development initiative itself) that is designed to attain specific policy objectives. That initiative is just one of many alternative ways that might be formulated to achieve the policy goals. Additionally, it may be just one of several complementary initiatives that the organization is engaged in at any given time. Understanding these two levels and creating a display that locates the eventual policy goals downstream from the immediate goals of the capacity development initiative (e.g., Figure 1) will help the various stakeholders — donors, partners, and beneficiaries — to hold a more informed discourse, to appreciate their different concerns, and to set more realistic, achievable goals and expectations (e.g., more upstream or more downstream).

Construct an adequately detailed logical framework for the intervention. Most program planners and evaluators now agree on the utility of developing what is variously labeled a "program theory" (Rossi, 1987), a "theory of action" (Patton, 1997), or a "results chain" (Mayne, 2001). If no analysis of the logic and assumptions of the intervention exists, engaging the stakeholders in constructing one can be an important first step toward developing a common language and common points of reference to evaluate the effort. Even when a formal logic analysis exists, it is wise for evaluators to check its currency and completeness with the stakeholders.

Explore and acknowledge the existence of factors and events in the environment of the project that affect performance. Organizations exist within complex political, institutional, and social environments that can affect the capacity development initiative in overt or subtle ways. Early identification and formal acknowledgment of these factors and their effects reduces the danger of exaggerating the role that the capacity development effort has played in the evolution of the organization. Conscious advantage may be taken of these factors. For example, in the case of the PM&E project, stakeholder awareness of the need for organizational change was a powerful force for mobilizing support for implementing new management systems.

Lessons Associated with Evaluation of the Capacity Development Initiative

Develop and agree upon conceptual frameworks to capture complex concepts and ensure their common understanding. Capacity development initiatives present a series of complex challenges for evaluators. Evaluation efforts may not be understood in the same way by all of the stakeholder groups who are engaged in them and supposed to be working in concert. They can help stakeholders to start by capturing in a display format simple concepts such as the impact chain, the logic model representing the effort, and a framework for organizational analysis (see for example Figures 1, 2, and 3). Once the evaluation team and the stakeholders have a comfortable common understanding of these, progress toward building a more integrated evaluation framework can be made (see for example Figure4). These displays and frameworks represent the conceptual and operational points of reference shared by the evaluation team and the stakeholders. Without them it can be difficult to ensure a constructive discourse among donors, partner agencies, beneficiaries, and the evaluation team.

Maintain clarity and consistency surrounding the level(s) of analysis (individual, department/program, organization) required to con-

firm project effects. Stakeholders have a tendency to assume that enhanced individual knowledge and skills can be taken as sufficient evidence of organizational capacity development. This tendency may arise from the desire to avoid complexity. However, although resources — including human resources — are the source of an organization's capabilities, organizational capacity is the outcome of deploying "bundles" of resources in particular ways in order to achieve strategic goals. It may therefore be appropriate to use individuals as the first level of analysis to determine that the requisite knowledge and skills are present, but it will also be necessary to use larger units of analysis to determine if the desired capabilities have been institutionalized and can be deployed when required to create effective work routines that accomplish complex tasks in pursuit of strategic goals. Capabilities that cannot be deployed effectively when required or are lost when individuals leave the organization cannot be regarded as "organizational" capabilities.

Strive to distinguish means and ends. Without an adequate understanding of the role of the capacity development initiative within the context of broader development processes and goals, there can be a tendency to confuse means and processes with ends and products. For example, the accomplishment of complex tasks may require the deployment and mobilization of a large number of people and the effective integration of their various skills. Once such levels of teamwork have been reached, those stakeholders closest to this type of effort may begin to view teamwork as an end in itself as opposed to the means by which complex tasks are accomplished in order to achieve the strategic goals of the organization that in turn demonstrate improved performance.

Consider the wisdom of employing terms other than "impact assessment" or "impact evaluation" when describing and evaluating capacity development initiatives. The term "impact" has a unidirectional, authoritarian connotation that does not capture the character of initiatives that are fundamentally social and organizational in nature. For a partner or a donor organization to claim credit for the "impact" of a capacity development effort on the performance of a beneficiary organization is to exaggerate the partner or donor's contribution by ignoring other factors at play. It is also to court the wrath of the beneficiary organization, which knows that in the long run, it alone is responsible for achieving its goals and accomplishing its mission. Mayne (2001) persuasively suggests that complex attribution questions such as those involved in the outcomes brought about by public administration programs can be handled effectively

by "contribution analysis." His employment of the results chain as the principal tool of contribution analysis in a six-step process could equally well be applied to assessing the contribution of efforts to develop organizational capacity in beneficiary organizations of the kind described in this article.

CONCLUSION

Capacity is the potential possessed by an organization for engaging resources and skills in optimum combinations so that it can perform relevant activities in order to accomplish its mission. Resources and skills are not usually productive on their own, and so organizational capacity development cannot be reduced to the simple delivery or acquisition of resources. Capacity development may include the acquisition of resources, but it must also include learning how to deploy and integrate these resources to accomplish complex tasks in line with its goals and strategy. It is the complexity of this inevitably social and political process of organizational change, in which both internal and external agencies are usually involved, that presents the challenge to evaluation. Input-output models cannot capture the intricate processes involved in the capacity development of complex systems. If efforts to evaluate capacity development are to provide useful feedback to donors, beneficiaries, and change agents to guide future practice, they must involve these stakeholders from the beginning. The evaluator can help to enhance the rigor of such participatory evaluations by striving to create clear frames of reference including agreed upon indicators and valid models for analysis drawn from organizational theory that are understood and endorsed by all stakeholders.

ACKNOWLEDGEMENTS

The authors acknowledge the financial or in-kind assistance provided by the Australian Centre for International Agricultural Research (ACIAR); Concordia University; the Danish International Development Agency (DANIDA); the Dutch Ministry of Foreign Affairs; the International Development Research Centre (IDRC); the International Fund for Agricultural Development (IFAD); the International Service for National Agricultural Research (ISNAR); and the Swiss Agency for Development and Cooperation (SDC).

REFERENCES

Alston, J.M., Norton, G.W., & Pardey, P.G. (1995). *Science under scarcity: Principles and practice for agricultural research evaluation and priority setting*. Ithaca, NY: Cornell University Press in cooperation with the International Service for National Agricultural Research.

Amit, R., & Schoemaker, P. (1993). Strategic assets and organizational rents. *Strategic Management Journal, 14*, 33–46.

Anderson, J.R., & Dalrymple, D.G. (1999). *The World Bank, the Grant Program, and the CGIAR: A retrospective view* (OED Working Paper series No. 1). Washington, DC: Operations Evaluation Department, The World Bank.

Austin, C. (1993). *The process of change: A synthesis study of "Institutional Capacity Development" projects for ODA*. London: Overseas Development Administration.

Berk, R.A., & Rossi, P.H. (1990). *Thinking about program evaluation*. Newbury Park, CA: Sage.

Bickman, L., & Rog, D. (Eds.). (1998). *Handbook of applied social research methods*. London: Sage.

Biggs, S.D. (1990). A multiple source of innovation model of agricultural research and technology promotion. *World Development, 18*(11), 1481–1499.

Bojanic, A., Hareau, G., Posada, R., Ruíz, A.M., & Solís, E. (1995). *Monitoring*. The Hague: International Service for National Agricultural Research (ISNAR).

Cheaz, J., de Souza Silva, J., Andersen, A., & Horton, D. (1999). *Introduction to the PM&E Project: Objectives, strategies, activities and outputs*. The Hague: International Service for National Agricultural Research (ISNAR).

CTA. (1998). *Assessing the impact of information and communication management on institutional performance.* Proceedings of a CTA workshop, Wageningen, January, 1998. Wageningen, The Netherlands: Technical Centre for Agricultural and Rural Cooperation (CTA).

Gálvez, S., Novoa, A.R., de Souza Silva, J., & Villegas, M. (1995). *The strategic approach to agricultural research management*. The Hague: International Service for National Agricultural Research (ISNAR).

Gijsbers, G., Janssen, W., Hambly Odame, H., & Meijerink, G. (Eds.). (2001). *Planning agricultural research: A sourcebook*. Wallingford, UK: CAB International.

Goyder, J. (1985). Face-to-face interviews and mailed questionnaires: The net difference in response rate. *Public Opinion Quarterly, 49*, 234–252.

Granger, A., Grierson, J., Quirino, T.R., & Romano, L. (1995). *Evaluation*. The Hague: International Service for National Agricultural Research (ISNAR).

Grant, R.M. (1995). *Contemporary strategy analysis: Concepts, techniques, applications* (2nd ed.). Oxford: Blackwell.

Guba, E.G., & Lincoln, Y.S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

Hilderbrand, M.E., & Grindle, M.S. (1995). Building sustainable capacity in the public sector: What can be done? *Public Administration and Development, 15*, 441–463.

Horton, D. (1986). Assessing the impact of international agricultural research and development programs. *World Development, 14*(4), 453–468.

Horton, D. (Ed.). (2001). *Learning about capacity development through evaluation: Perspectives and observations from a collaborative network of national and international organizations and donor agencies*. The Hague: International Service for National Agricultural Research (ISNAR).

Horton, D., Ballantyne, P., Peterson, W., Uribe, B., Gapasin, D., & Sheridan, K. (1993). *Monitoring and evaluating agricultural research: A sourcebook*. Wallingford, UK: CAB International.

Horton, D., Ballantyne, P., Peterson, W., Uribe, B., Gapasin, D., & Sheridan, K. (1994). *Seguimiento y evaluación de la investigación agropecuaria: Manual de referencia*. Wallingford, UK: CAB International.

Horton, D., & Mackay, R. (1998). Assessing the organizational impact of development cooperation: A case from agricultural R&D. *Canadian Journal of Program Evaluation, 13*(2), 1–28.

Horton, D., Mackay, R., Andersen, A., & Dupliech, L. (2000). *Evaluating capacity development in planning, monitoring and evaluation* (ISNAR Research Report No. 17). The Hague: International Service for National Agricultural Research (ISNAR).

Hudson, M. (1999). *Managing without profit: The art of managing third-sector organizations* (2nd ed.). London: Penguin.

Jackson, E.T., & Kassam, Y. (1998). *Knowledge shared: Participatory evaluation in development cooperation*. Ottawa: International Development Research Centre.

Kirkpatrick, D.L. (1998). *Evaluating training programs: The four levels*. San Francisco, CA: Berrett-Koehler.

Knox, C. (1996). Political context and program evaluation: The inextricable link. *Canadian Journal of Program Evaluation, 11*(1), 1–20.

Lusthaus, C., Anderson, G., & Murphy, E. (1995). *Institutional assessment: A framework for strengthening organizational capacity for IDRC's research partners*. Ottawa: International Development Research Centre.

Maconick, R., & Morgan, P. (1999). *Capacity building supported by the United Nations: Some evaluations and some lessons*. New York: United Nations.

Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation, 16*(1), 1–24.

Montague, S. (1997). *The three Rs of performance: Core concepts for planning, measurement, and management*. Ottawa: Performance Management Network Inc.

Moore, M. (1995). *Institution building as a development assistance method: A review of literature and ideas*. Stockholm: Swedish International Development Authority.

Murphy, J. (2000). Epilogue. In D. Horton, R. Mackay, A. Andersen, & L. Dupliech (Eds.), *Evaluating capacity development in planning, monitoring and evaluation* (pp. 79–86). The Hague: International Service for National Agricultural Research (ISNAR).

Novoa, A.R., & Horton, D. (Eds.). (1994). *Administración de la investigación agropecuaria: Experiencia en las américas*. Santa Fe de Bogotá: Tercer Mundo editores.

Novoa, A.R., & Horton, D. (1999). *Dinámica de la PS&E de la investigación agropecuaria en América Latina y el Caribe: Un informe sobre nueve estudios de caso* (part of the series, Evaluating capacity development in planning, monitoring, and evaluation). The Hague: International Service for National Agricultural Research (ISNAR).

Patton, M.Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.

Peterson, W., & Perrault, P. (1998). Agricultural research organizations: The assessment and improvement of performance. *Knowledge, Technology and Policy, 11*(1&2), 145–166.

Prahalad, C.K., & Hamel, G. (1990, May–June). The core competence of the corporation. *Harvard Business Review*, 71–79.

Rist, R.C. (1995). Postscript: Development questions and evaluation answers. In R. Picciotto & R.C. Rist (Eds.), *Evaluating country development policies and programs: New approaches for a new agenda* (New Directions for Evaluation No. 67) (pp. 167–174). San Francisco, CA: Jossey-Bass.

Röling, N.G., & Wagenmakers, M.A.E. (Eds.). (1998). *Facilitating sustainable agriculture*. Cambridge: Cambridge University Press.

Rossi, P.H. (1987). The iron law of evaluation and other metallic rules. *Research in Social Problems and Public Policy, 4*, 3–20.

Rossi, P.H., Freeman, H.E., & Lipsey, M. (1999). *Evaluation: A systematic approach* (6th ed.). Newbury Park, CA: Sage.

Rowe, A. (Ed.). (1999). Empowerment evaluation [Special issue]. *Canadian Journal of Program Evaluation, 14*.

Taschereau, S. (1998). *Evaluating the impact of training and institutional development programs: A collaborative approach* (EDI Learning Resource Series). Washington, DC: World Bank.

Yin, R.K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage.