ILAC Working Paper 2

Evaluation, Learning and Change in Research and Development Organizations: Concepts, Experiences, and Implications for the CGIAR

Douglas Horton, Viviana Galleno, and Ronald Mackay

2nd Edition

November 2007

Institutional Learning and Change (ILAC) Initiative - c/o Bioversity International
Via dei Tre Denari 472°, 00057 Maccarese (Fiumicino), Rome, Italy
Tel: (39) 0661181, Fax: (39) 0661979661, email: **ilac@cgiar.org**, URL: **www.cgiar-ilac.org**

The ILAC initiative fosters learning from experience and use of the lessons learned to improve the design and implementation of agricultural research and development programs. The mission of the ILAC Initiative is to develop, field test and introduce methods and tools that promote organizational learning and institutional change in CGIAR centres and their partners, to expand the contributions of agricultural research to achievement of the Millennium Development Goals.

This paper has been reformatted to comply with the style of the ILAC Working Paper series.

Citation: Horton, D., Galleno, V. and Mackay, R. (2007) (2nd Edition) *Evaluation, Learning and Change in Research and Development Organizations: Concepts, Experiences, and Implications for the CGIAR.* ILAC Working Paper 2, Rome, Institutional Learning and Change Initiative.

Originally published by the International Service for National Agricultural Research as: Horton, D., Galleno, V. and Mackay, R. (2003) *Evaluation, Learning and Change in Research and Development Organizations: Concepts, Experiences, and Implications for the CGIAR.* Discussion Paper 3. The Hague: International Service for National Agricultural Research.

Table of Contents

Evaluation, Learning and Change in Research and Development Organizations: Concepts, Experiences, and Implications for the CGIAR

Douglas Horton, Viviana Galleno and Ronald Mackay

November 2007

Abstract

The Consultative Group for International Agricultural Research (CGIAR) is under strong pressure to enhance its capacity to learn and change. Business as usual is no longer an option in an environment that demands the increasing effectiveness, efficiency and relevance of the services and products provided by its centers. The direction CGIAR centers and the system need to pursue for learning and change will require the right blend of central guidance (e.g. from the Science Council and the Executive Committee) and center-led self-evaluation and change initiatives. Over time, the role of the external bodies and external program and management reviews (EPMR) can become focused on the integrity and quality of internal evaluation and quality assurance systems within individual centers. The success of system governance and management mechanisms will ultimately rest on the degree to which they help build and complement the centers' own internal evaluation systems and strengthen the capacity of the centers to become learning organizations employing sound self-evaluation and self-improvement practices. Some practical suggestions for strengthening institutional learning and change in the CGIAR are offered.

Keyword: Learning, Change, CGIAR, R&D, Evaluation.

1. Introduction

Background

After two decades of expansion and success associated with the Green Revolution, since the early 1990s the CGIAR has been challenged to adapt itself to rapidly changing social, economic, political, technological and institutional conditions. The CGIAR has broadened its mandate beyond the generation and transfer of technology to expand food production, and now embraces the objectives of poverty reduction and protection of the natural environment. CGIAR centers that were established to generate production technologies have found it necessary to rethink their basic goals, structures, staffing, inter-institutional relationships, and management practices.

Demands for change have come during a period of declining support for international development assistance, particularly in the area of agriculture. CGIAR centers have found it increasingly difficult to obtain resources to carry out their work, and the proportion of "core" funding has declined to less than half of the total budget of the CGIAR, as donors target more of their contributions to specific projects or topics.

These pressures for change have been accompanied by growing demands for accountability and evidence of impact as well as institutional change with a view to enhancing CGIAR performance.

Reflecting concerns over the effectiveness and utility of evaluation in the CGIAR, in February 2002, the International Center for Maize and Wheat (CIMMYT) and the Standing Panel on Impact Assessment (SPIA) organized an international conference on impact assessment to

address the basic question: "Why has impact assessment research not made more of a difference?" At the conference, attention was drawn to alternative ways of conducting evaluations and providing information about the results of agricultural research to stakeholders for a broader variety of purposes than the promotion of public relations. In particular, two kinds of evaluation were discussed, one that assesses the economic effects of research and another that helps organizations learn, change and improve their performance.

At the CIMMYT-SPIA conference in Costa Rica in 2002, interest emerged in further discussing the use of evaluation for learning and change. As a result, a workshop on "Institutional Learning and Change" was held at the International Food and Policy Research Institute (IFPRI), in Washington D.C., on February 4-6, 2003.

It is a positive sign that key stakeholders wish to go beyond the limitations of impact assessment and address the challenge of improving the CGIAR's activities and performance through organizational learning and change.

Purpose and organization of the paper

Our aim in this paper is to contribute to the on-going discussion of evaluation, learning and change in the CGIAR. Our basic theses are as follows:

- Individual and organizational learning are essential to promote institutional changes that bring about significant improvements in to the performance of CGIAR centers and the system as a whole.
- Evaluation can play a valuable role in promoting the needed learning and changes.
- For evaluation to play this role, a more reflective, evaluative, learning-oriented culture needs to be nurtured in the CGIAR.

The paper has four main sections. After this introduction, in Section 2 we review experiences with evaluation, learning and change in the CGIAR. In Section 3, we introduce some basic concepts and definitions and summarize relevant experiences in the fields of evaluation, organizational learning, and change. In Section 4, we draw some general conclusions from the foregoing analysis and present some implications for strengthening evaluation, organizational learning, and change in the CGIAR.

Our primary purpose is to draw attention to the contribution that evaluation can make to promoting organizational learning and inducing institutional change in the interest of improved performance in the areas to which the centers of the CGIAR have committed themselves. Experience in both the public and private sectors shows that evaluation engaged in as sincere, participative enquiry can provide the information needed by members of Boards of Trustees, senior and middle managers, their teams and individual staff members to learn in the face of adversity, to change and to guide their organizations towards greater success.

Institutions and organizations

The terms organization and institution are often confused. In this paper the term institution refers to a socially sanctioned and maintained set of established practices, norms, behaviors, or relationships (e.g., trade regulations, land tenure systems, and family) that persist over time in support of collectively valued purposes. Organizations are formal structures with designated roles and purposes (Morgan and Qualman, 1996: 1). According to Burki and Perry (1998: 11),

> Institutions are formal and informal rules and their enforcement mechanisms that shape the behavior of individuals and organizations in society. By contrast, organizations are entities composed of people who act collectively in pursuit of shared objectives. These organizations and individuals pursue their interests within an

> institutional structure defined by formal rules (constitutions, laws, regulations, contracts) and informal rules (ethics, trust, religious precepts, and other implicit codes of conduct). Organizations, in turn, have internal rules (i.e. institutions) to deal with personnel, budgets, procurement, and reporting procedures, which constrain the behavior of their members.

Uphoff (1986) distinguishes between organizations that are not institutions (e.g., a recently established agricultural consultancy firm), institutions that are not organizations (e.g., marriage), and organizations that are also institutions (e.g., a country's central bank or its national agricultural research institute).

2. Experiences with Evaluation, Learning and Change in the CGIAR

International agricultural research centers and the system's governing bodies have carried out reviews and evaluations since the CGIAR was established in the early 1970s. Under the coordination of the interim Science Council (iSC) and the CGIAR Secretariat, centers are reviewed approximately each five years by external experts, through EPMRs.[2] Centers also carry out periodic internal program reviews (IPRs) and occasional center-commissioned external reviews (CCERs). Some assess the economic impact of their more successful programs and technologies (Technical Advisory Committee (TAC), 2001).

Despite this range of evaluative activities, in the mid-1990's, donors requested improvements in the CGIAR's assessment of its performance and impact. The response was the establishment of an independent evaluation function reporting to the CGIAR as a whole (CGIAR, 1995: 11). The Impact Assessment and Evaluation Group (IAEG) was intended to spearhead impact assessment at the system level. Later, this body evolved into the Standing Panel on Impact Assessment (SPIA), within the interim Science Council (iSC) of the CGIAR.

Since 1995, the number and quality of economic impact assessments carried out by CGIAR centers and at the system level have increased. Measures have also been taken to improve the conduct and rigor of the EPMRs (TAC, 1997). Yet, many of the CGIAR's donors and other stakeholders still feel that evaluations in the CGIAR are not producing the organizational learning and institutional changes needed for the centers to retain or reestablish their positions as leaders in the field of agricultural research and development, and to make more significant contributions to the broad goals of poverty alleviation, environmental protection and food security.

Organizational learning, a major theme in the fields of management and organizational development in private firms, foundations, and non-governmental organizations, has made little headway in the CGIAR. There has been considerable pressure for large-scale organizational change over the last decade, but the CGIAR response has often been to focus on structural changes. The Ford Foundation has supported an ambitious Organizational Change Program that seeks to strengthen leadership and organizational performance by supporting innovative ways of managing collaborative alliances, and improving knowledge management. Over the last five years, the Program's activities have included retreats and workshops for senior managers and center staff on leading organizational change, resource

---

[2] Originally, the CGIAR implemented external program reviews. In the early 1980s, external management reviews were added. Later in the 1980s, the program and management reviews were merged into a single External Program and Management Review (EPMR). Recently, there have been moves to separate the program and management reviews once again.

mobilization, and promoting effective teamwork and collaborative partnerships. The Program has also sought to develop knowledge management strategies at center level and a working group exists to support the system's efforts in information and communication technology and knowledge management (Spink, et al., 2000). However, there has been little effort so far to employ evaluation systematically to promote organizational change and performance improvement.

Despite efforts to integrate these various activities and link them to decision making, there is as yet no integrated evaluation system or planning, monitoring and evaluation system within the CGIAR.

While evaluation is widely used as a device to promote organizational learning in much of the private and public sector, within the CGIAR there has been a tendency to reserve the term for measuring the economic effects of research (Alston, Norton and Pardey, 1995). There is some evidence that this may be changing, however. There have been recent calls for a broadening of the agenda of impact research so as to address the concerns of, and be accessible to, a wider array of audiences than just the readers of professional economics journals e.g. taxpayers, development agencies and philanthropic organizations as well as researchers and the organizations within which they work. (Maredia, Byerlee, and Anderson, 2000; Pingali, 2001).

These calls imply that impact assessments should serve the needs of two broad groups of users simultaneously. One user group represents the traditional clients of impact assessment i.e. those "external" to research organizations, such as governments and donors who require accountability for the funds they invest, the other group represents clients "internal" to research centers, such as boards of trustees, managers, stakeholder advisory committees and scientists who need information to help improve their oversight and management practices.

Experience suggests that it is a mistake to assume that the same studies can serve such diverse interests. In cases where the needs of these two user groups present a dilemma for the impact assessor, the organization, its managers and other decision makers tend to be the casualty. Substantially less emphasis has been placed on using impact assessment to serve the information and learning needs of the organization than on addressing the accountability needs of external clients.

Moreover, the language used to describe more learning-oriented studies suggests that these are considered to be an inferior sub-category of impact assessment. They are referred to as "partial impact studies" – as opposed to "more comprehensive impact studies" whose "lack of rigor often undermines their credibility" (Maredia, Byerlee, and Anderson 2000 p.6).

Much f the controversy surrounding the current and potential use of impact assessment stems from the dominance of an arbitrary definition of evaluation within the CGIAR. It would appear that what SPIA and center economists do is "economic evaluation". Equating a specific method (e.g. impact assessment) with a field of inquiry (e.g. evaluation) is an example of what has been called the "law of the instrument" (Kaplan; 1964). By invoking, unintentionally, the "law of the instrument', the CGIAR has unilaterally defined evaluation to mean impact assessment to the exclusion of other evaluative activities that, within the broader external evaluation community, are also generally included within the term "evaluation".

There are signs that SPIA is beginning to be aware of this anomaly. Recently, it has noted that "Different types of assessment and evaluation have different functions and should be executed by different actors in the System. Impact assessment should not be confused with programme evaluation" (iSC Secretariat, 2002). In addition, the work of SCOER/SPIA (Standing Committee on External Reviews/Standing Panel on Impact Assessment) working group suggests steering the CGIAR away from overly narrow thinking towards an appreciation of the range of evaluation practices that exist and the uses to which they can be put (SDR/TAC:

IAR/01/04). This line of thinking appears to parallel that of the European Initiative for Agricultural Research and Development (EIARD, 2002). Nevertheless, it is likely to take some time for the full range of evaluation approaches and practical evaluation methods, for example, those described in Horton et al., 1993, to be embraced and mainstreamed.

The CGIAR's use of idiosyncratic terminology that equates evaluation with impact assessment has unfortunate implications. One is that evaluators within the CGIAR with roots in the broader evaluation community tend to be excluded from "evaluation" activities. When interaction has occurred, the two "camps" have tended to talk a different language. A second implication is that future roles for the expansion of evaluation to serve institutional and learning needs is made doubly difficult. The traditional public relations function of impact assessment gets in the way of thinking about evaluation as a means of pursuing institutional learning and change.

For the CGIAR to make better use of evaluation as a tool for organizational learning and performance improvement, two key issues need to be resolved:

- Can impact assessment be recast so as to meet the information needs of diverse stakeholders or must other evaluation types be embraced to satisfy different requirements?
- Can qualitative data and mixed methods be admitted into evaluation studies or are only quantitative data admissible?

How these two questions are answered will depend upon the CGIAR's willingness to embrace and experiment with broader theoretical orientations and modes of knowledge production currently employed in a wide range of public and private organizations.

3. Experiences in the Broader Fields of Management and Evaluation

Evaluation: meanings, methods, and controversies

Definitions of evaluation

The term "evaluation" is defined and used in a variety of ways. Webster's New World Dictionary defines evaluation broadly as judging or determining the worth, value, or quality of an entity. The Joint Committee on Standards for Educational Evaluation in the United States of America (Sanders, 1994) defines evaluation as the systematic investigation of the worth or merit of an object (such as a project, program or policy). Carol Weiss (1998: 4) defines evaluation as the systematic assessment of the operation and/or the outcomes of a program or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy.

In the USA, the field of evaluation as a transdiscipline and profession originated in the 1960s, when federal legislation required the evaluation of social programs. For this reason, the term "program evaluation" is common there. Program evaluation has been defined as "the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming" (Patton, 1997). The role of evaluation in diagnosing organizational weaknesses and contributing to organizational learning and change is growing (Patton, 1999; Preskill and Torres, 1999).

Evaluations are required for many international development programs, and a specialized branch of evaluation has grown up to satisfy this need (Picciotto and Rist, 1995; Picciotto and Wiesner, 1998). In this context, the Development Assistance Committee of the Organisation

for Economic Co-operation and Development (OECD, 1991) defines evaluation as "an assessment, as systematic and objective as possible, of an on-going or completed project, program or policy, its design, implementation and results."

The fact that different groups define evaluation differently can lead to conflict over what evaluation *is*, *how* it should be done, and what constitutes *good* evaluation.

The CGIAR provides an illuminating example of the controversies surrounding what constitutes "good evaluation." In 1997, the IAEG commissioned a metaevaluation of impact studies carried out by CGIAR centers. Dr. Leslie Cooksy, a respected professional in the USA program evaluation community, was selected to carry out the study and all CG centers were requested to send her copies of their impact studies. Cooksy established a set of standards for evaluating the impact studies. These included:

1. Claims made by each study
2. Products reported on and uptake of these
3. Effect size, nature and direction of uptakes and outcomes
4. Source of evidence to support the statement of effect
5. Strength of support provided
6. Plausibility of the effect reported
7. Concerns about the sources of evidence
8. Alternative explanations for the outcome reported
9. Appropriateness of any generalization made to a higher level of aggregation

In an early stage of the metaevaluation, a large number of the studies received were found not to represent impact evaluation efforts and discounted. Later, a panel of expert judges found that 76 out of the 87 studies retained did not meet the criteria established for an impact study. While these 76 studies contained good coverage of CGIAR activities and effects, the discussions of methodology tended to be weak. Only 11 studies were considered to approximate to being exemplars of good evaluation practice. One of these was a study of the International Service for National Agricultural Research (ISNAR) impact, carried out in 1996 (Mackay et al., 1998). Cooksy recommended that the following strategies, some of which had been modeled in each of the 11 studies, would strengthen future CGIAR evaluations:

- A clear description and rationale for the sampling of cases, respondents or documents reviewed
- Use and synthesis of evidence across multiple sources
- Disclosure of data gaps and limitations, and cautious reporting when faced with severe data gaps and limitations
- The use of a logical framework to organize the information presented.

As might be expected, the results of the metaevaluation were not well received by those who had carried out the studies in question. At the annual meeting of the CGIAR in 1997, Jock Anderson, a leading agricultural economist with the World Bank and a champion of economic impact assessment, presented a stinging critique of the metaevaluation which seriously questioned its approach, methods and results. Consequently, the IAEG never formally issued the report.

As a postscript, in 2002, John Gardner, an agricultural economist at the University of Maryland, was commissioned by the World Bank to conduct another metaevaluation of CGIAR impact studies. The same ISNAR study that Cooksy had considered to bean exemplar of good evaluation practice was considered by Gardner to be of little relevance or use.

Evaluation purposes

Evaluations are carried out for various reasons and fall under three main headings: accountability, decision-making, and knowledge.

Evaluations are often carried out *to satisfy accountability requirements*. Those who fund research or development projects need assessment of the extent to which their objectives have been achieved and their resources have been well used. This is the main reason for the hundreds – perhaps thousands – of end-of-project reviews and evaluations that are carried out each year for donor-funded projects in developing regions. Most of these accountability-oriented evaluations, focused on goal attainment and resource use, are carried out by external evaluators in compliance with donors' requirements. In the CGIAR, EPMRs are the principal accountability mechanism for core-funded center activities. Many donors also evaluate projects they fund in CGIAR centers. Accountability-oriented evaluations generally produce reports which are delivered to the donor or their representative (such as the Science Council and Executive Committee of the CGIAR) for their information and action.

The second purpose of evaluation is *to provide information that managers or policy makers can use* to improve projects, programs or policies. Decision makers may engage individuals from inside the program or organization or from outside, to carry out the evaluation and report on findings. In such cases, formal evaluation reports are less frequently produced and distributed. Sometimes reports remain confidential or in draft form. Face-to-face communication of findings is often more important.

A third reason for carrying out an evaluation is *to generate new knowledge or learn lessons* about a project, a program, a policy or an organization. Such information and insights may be put to use by different groups to improve the design, implementation or evaluation of future activities. Evaluations carried out for this purpose range from participatory exercises designed to draw lessons for the management of a project team or an organization (Estrella, 2000) to quantitative studies employing quasi-experimental methods that aim to rigorously test hypotheses (Alston et al, 1995).

Exhibit 1. Three Basic Purposes of Evaluation

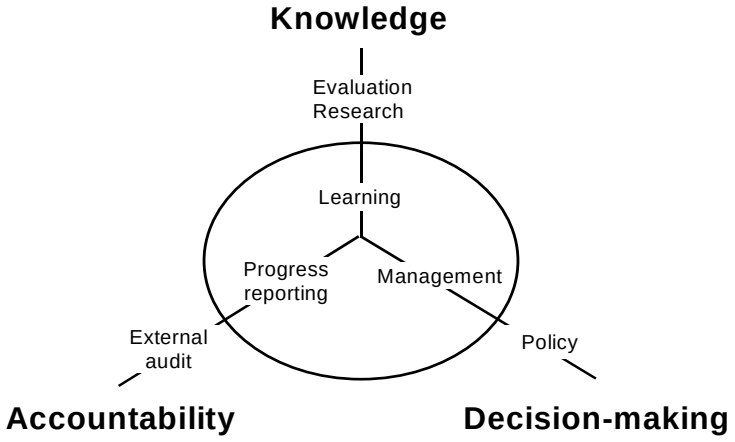
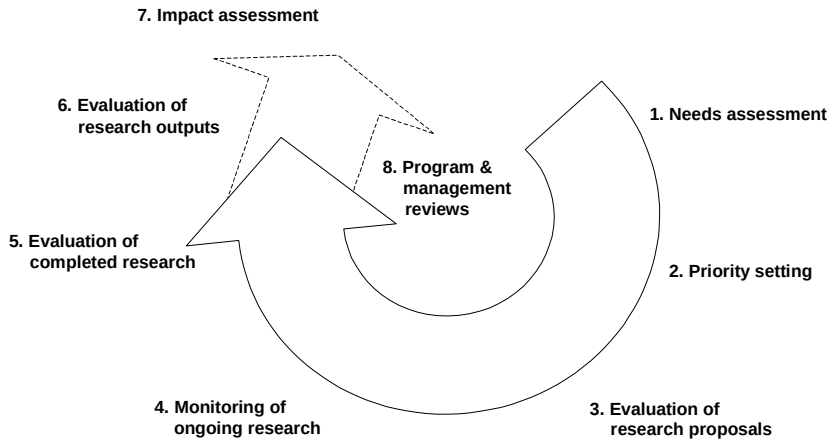


Exhibit 2. Eight Types of Evaluation



Source: Horton, 1997.

Terms of reference for evaluations often include elements of each of these three different purposes. For example, the above-cited OECD document indicates that the aim of an evaluation is "to determine the relevance and fulfillment of objectives, developmental efficiency, effectiveness, impact and sustainability. As such, an evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors."

However, in practice it is difficult to combine such diverse purposes in a single evaluation that will satisfy the information needs of diverse stakeholder groups. An evaluation that does a good job of satisfying external accountability requirements may be of little use to managers or program staff. And an evaluation that provides managers with useful information for improving on-going work may be of little use to donors, who need to report on results to their taxpayers.

Exhibit 2 illustrates eight distinct types of evaluations that are carried out in agricultural research organizations, and how they relate to the stage of research at which the evaluation takes place. At the risk of over-simplification, it can be said that the first four types – needs assessment, priority setting, evaluation of research proposals and monitoring of on-going research – are usually carried out to support internal decision making, and the next four types – evaluation of completed projects, evaluation of research outputs, impact assessment and program reviews – are more often carried out to meet external accountability requirements. Only two types of agricultural research evaluation have generated new knowledge that has been captured and disseminated in publications: priority setting and economic impact assessment. There is little evidence that lesson learning has featured prominently in these or the other types of evaluation in agricultural research organizations.

Evaluation approaches and methods

Different types of evaluation call for different evaluation methods. Exhibit 3 presents a summary of the methods commonly associated with different types of evaluation carried out in agricultural research organizations.

Exhibit 3. Types and Methods of Agricultural Research Evaluation

Evaluation type	Evaluation methods
1. Needs assessment	– Expert judgment – Literature review – Farmer, market & consumer surveys – Macro-economic studies
2. Priority setting	– Expert judgment – Ex-ante assessment of economic impact – Social or environmental impact assessment – Scoring methods – Participatory planning
3. Review of research proposals	– Peer review – Expert review – Stakeholder review
4. Monitoring of on-going research	– Internal reviews (by peers) – External reviews (by peers/experts)
5. Evaluation of completed research	– Project completion reports – External evaluation by peers/experts
6. Evaluation of research outputs	– Peer review of publication – Bibliometric studies – On-farm trials / participatory evaluation of technology – Producer, market & consumer surveys
7. Impact assessment	– Adoption studies – Economic evaluation – Social & environmental assessment methods
8. Program and management reviews	– Internal reviews (by peers) – External reviews (by peers/experts) – Internally commissioned external reviews (by peers/experts) – Financial audit – Comprehensive audit

The idea of a "single best method" approach is inherent in the admonition that "research evaluation and priority-setting methods should attempt to use measures that approximate changes in economic surplus" (Alston et al., 1995 p.503). On the other hand, a move towards multiple methods is given voice by Patton (op. cit. 2002 p.105) when he says that a more utilization-oriented evaluation advocates "presenting primary intended users with a range of options, facilitating their understanding of the strengths and weaknesses of those options, and then helping them determine the "best fit" for their needs.

Evaluation controversies

The field of evaluation, like most relatively "young" fields of study and practice, is not without its controversies. In order to understand some of the major controversies it is necessary to understand some basic terms used in the philosophy of science. The history of the philosophy of science documents man's pursuit of a few fundamental and basic questions:

What is the world? How can we come to know it? How can we control it for our own purposes? (Lincoln and Guba, 1985, p.14).

Despite centuries of thought and writing, there is no universally-accepted answer to these questions. *Epistemology* is the theory of how we come to acquire knowledge of the external world. The term is often used to refer to the methods or formal procedures which lead to the acquisition of knowledge. Different *paradigms* or schools of thought determine what methods or procedures we can employ. For example, a theory founded on positivism would lead to different methods from a theory founded on constructivism (Van der Veen, 2000).

Positivism is the paradigm upon which scientific method is based. It asserts that knowledge is based on observable, measurable natural phenomena and the relationships that exist between them. *Constructivism*, on the other hand, is a paradigm on which many of the methods employed in the organizational and management sciences – and disciplines such as anthropology – are based. It asserts that reality is constructed socially by people through individual and collective agreement on the perceived situation (Van der Veen).

Positivists would criticize constructivists for postulating unobservable structures such as "organizational culture". Constructivists would accuse positivists of treating only the superficial, observable level of social life and ignoring the complex mechanisms that generate social phenomena (Abercrombie, et al., 1994; Berger et al., 1967; Worthen, et al., 1997).

Different disciplines are wedded, to a greater or lesser extent, to one or other of these two paradigms and therefore determine how those brought up within a given discipline are permitted to conduct inquiry activities.

To a large extent, the controversies that flourish in evaluation arise out of the disciplinary paradigm or tradition of evaluation practitioners. Their disciplinary training will predispose them to display different philosophical and ideological beliefs, methodological preferences, and practical choices based on prior experience.

While the philosophical and ideological belief in positivism dominated evaluation for decades, today, there is an appreciation that the choice between positivism and constructivism is not an "either or" issue. There is a belief that the epistemology drawn upon must be right for the evaluation being undertaken.

The different philosophical assumptions and positions have given rise to different methods of inquiry. *Quantitative methods* emphasize experimental design, control, instrumentation, measurement of objective data (usually outputs or outcomes), and statistical methods of analysis based on deduction and hypothesis-testing. In contrast, *qualitative methods* are usually conducted in natural settings, such as organizations. They use the evaluator as the principal instrument, employ multiple data-gathering methods (including observation, document review, and interviewing), emphasize "thick description", examine processes as well as outcomes, and use inductive approaches to data analysis to illuminate salient patterns of meaning from the perspective of those being studied.

Practical decisions about which evaluation "model" or approach to adopt in any particular evaluation can still present a difficult choice. In a recent review of evaluation approaches, Daniel Stufflebeam – one of the founders of modern evaluation – characterized 22 different approaches that emerged between 1960 and 1999, and assessed their strengths and weaknesses (Stufflebeam, 2001).

These controversies have been described as representing different extremes of an evaluation continuum with the natural science model of "pure research" at one end and the "utilization-focused" approach to evaluation at the other (Mackay and Horton, 2003). There has been a trend within the broader evaluation community for strict adherence to one or other of these extremes to give way to a view that appropriate evaluation requires "pluralism" (Sechrest, 1993), "mixed-methods" (Greene and Caracelli, 1997), and "responsiveness" (Greene and

Abma, 2001). Nevertheless, the controversy surrounding ideologies, methodologies and practical choices still flourishes on the sidelines (Rossi, 1994).

Controversies also exist, to some extent, within the CGIAR. Possibly because of the predominance of disciplines with a positivist foundation, evaluation within the CGIAR has tended to become almost synonymous with economic evaluation conducted using quantitative methods. While other kinds of evaluation have been carried out, some of them relatively frequently and routinely, they have either been given other names, such as "review" as in the EPMRs and internal program reviews, "management research", "social science research," or "partial impact assessment studies" (Maredia et al., 2000). It is only relatively recently that a more inclusive perspective is emerging, so that different kinds of evaluation with roots in different philosophical and ideological traditions are included in the term "evaluation" (Horton, 1998). Nevertheless, it may take some time for the full range of available evaluation approaches and practical evaluation methods to be embraced and mainstreamed in the CGIAR.

Organizational learning: meanings and processes

Modern organizations must adapt to changing circumstances and successful adaptation depends upon the encouragement people get to learn individually and in groups and then to translate their learning into improved work processes and practices. In this section, we introduce some of the key concepts relating to organizational learning.

Basic definitions

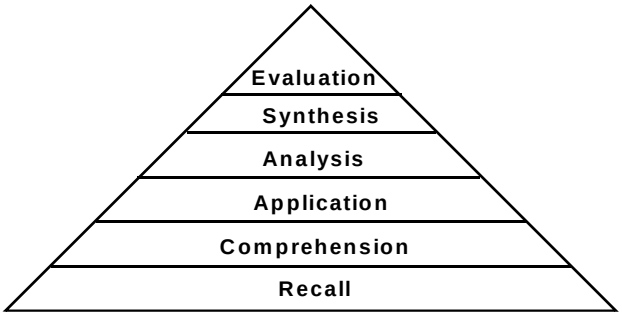
Simply put, *learning* is the acquisition of new knowledge that is relevant and pertinent to the learner. New knowledge is important because it offers the potential to bring about changes in the behavior of individuals, groups and organizations. The capacity to change on the basis of acquiring new, pertinent knowledge can be crucial to the survival of an organization in today's environment (Argyris and Schön, 1978; Schön, 1983; Mabey, et al. 1994; 1995).

In 1956, Benjamin Bloom classified six levels of learning into a hierarchy known as Bloom's Taxonomy. The six levels are: recall, comprehension, application, analysis, synthesis, and evaluation.

The cognitive demands increase as one moves through the taxonomic hierarchy, with evaluation making the greatest demand and drawing upon all of the cognitive processes in the levels beneath it.

O*rganizational learning* has been defined as, "an organization's capacity for accumulating knowledge from its own experiences, disseminating that knowledge to members throughout the organization (and not to a single individual or group within it), reflecting on it and using it as a basis on which to build planning and programming activities, to adapt and to cope with change" (Brown; 1998:62). Brown goes on to note that, "learning is far more than the collection and collation for data and the preparation of reports. It is a process of collective (and collaborative) cogitation and reflection on experience, and requires the inculcation of positive attitudes to learning and the development of analytical and cognitive capabilities at the institutional level. One might think in terms of a *predisposition or propensity to learn* [originally highlighted]." A learning organization as one that facilitates the learning of all its members and continuously transforms itself to maintain sound performance (Dale, 1994, 22).

Exhibit 4. Six Levels of Learning



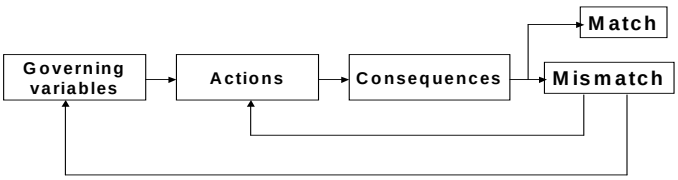
Source: Bloom, 1956.

Two kinds of organizational learning have long been distinguished within management studies, single-loop learning and double-loop learning (Exhibit 5) (Argyris, 1977, p.68 ff; Argyris, 1976). *Single-loop learning* occurs when an error is detected and corrected without questioning or altering the underlying values of the organization. Organizations that use only single-loop learning tend to focus on isolated, discrete and immediate problems without regard for the organization's fundamental goals or the rationale for how the organization chooses to do its work.

Double-loop learning occurs when errors are corrected by first examining the governing assumptions and then realigning the processes and activities accordingly. Organizations that employ double-loop learning embrace the creative thinking associated with challenging the organization as a complex goal-directed system. Double-loop learning involves individuals and groups acting together to uncover problems, create solutions and implement these in order to pursue and accomplish the mission driving the organization.

Knowledge has been defined in several ways. Maturana and Varela (1984) defined knowledge as effective action in the domain of existence. Such a definition emphasizes the importance of local, social and technical practices, competent performance and the interplay among practitioners (Leeuwis et al., 2002; Engel, 1997). Knowledge has also been described as justified belief that increases an entity's potential for effective action (Alavi, 2000). Information becomes knowledge when it is contextualized, given a valid meaning, seen as relevant, and capable of being related to practical experience. Knowledge increases the potential of an individual, group or organization to engage in effective action. It does not *necessarily* lead to effective action, but it does increase the potential.

Exhibit 5. Single-Loop and Double-Loop Learning



Source: Argyris, 1977.

Organizational knowledge is the knowledge that the organization possesses, that originates and is applied in the minds of the individual members of the organization. It becomes embedded in documents, databases and in the routines, processes and practices employed to get work done within the organization. Organizational knowledge is not a discrete or simple concept. Unless organizational knowledge is carefully managed, it often remains intuitive and therefore is hard to capture in words or to grasp completely in purely logical, easily explained terms (Davenport et al., 1997; Quintas et al., 1999).

A *learning organization* has been defined as one that consciously and intentionally promotes and facilitates the learning of all its members and uses the new knowledge created for its own self-improvement and enhanced performance (Pedler et al., 1991);. In learning organizations, "employees expand their capacity to create the results that they truly want, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free and where people are continually learning how to learn together" (Senge, 1994; 1990). A learning organization is able to capture and inter-relate the knowledge possessed by individuals and groups, make it explicit, and use it at the organizational level, on a continual basis (Jones and Hendry, 1992).

Learning is only one among several important knowledge processes that can be managed in organizations to help improve their performance. Collectively, these practices are known as *knowledge management*. Sharing knowledge, transforming knowledge to create innovative processes, services, and products, and transferring knowledge can all be managed well – or, as in most organizations – not at all.

The creation of expert systems also helps illustrates the difference between *managing knowledge* and the simpler process of *information management* (Open University B823, p.32). Information management stresses merely the delivery and accessibility of content and tends to assume that information capture can be standardized and automated. Managing knowledge, on the other hand emphasizes adding value to content by filtering, synthesizing, interpreting and adding context. There is a focus on organizational culture and work practices and this requires on-going human inputs whose variability precludes the capture process being reduced to a mere automated procedure.

Types of knowledge and modes of creating knowledge

Types of knowledge

How we categorize knowledge can have important implications for how we manage it and how it can be used to promote and maintain organizational learning and change.

Two kinds of knowledge exist within individuals and organizations – tacit knowledge and explicit knowledge (Nonaka and Takeuchi, 1995). *Tacit knowledge* is neither easily visible nor capable of being expressed simply. Tacit knowledge tends to be personal, context-specific, hard to formalize and challenging to communicate. For example, subjective insights, intuitions and hunches fall into this category.

Explicit knowledge has been codified so that it can be easily communicated and shared with others in the form of text, scientific formulae, procedures or principles.

For organizational learning, a key issue is the ease with which tacit and explicit knowledge can be shared. For successful organizational learning, a deliberate effort must be made to transform employees' tacit knowledge into explicit knowledge (Skyrme, 1999; 1997). Unless that happens, those with valuable tacit knowledge eventually leave the organization taking their knowledge with them and leaving the organization poorer.

The categories of tacit and explicit knowledge are not mutually exclusive. Tacit knowledge often exists prior to explicit knowledge. For example, expert systems, for use in agricultural extension work, are created by making explicit the tacit knowledge of successful farmers and experienced extension workers. This is done by eliciting it in words, writing it down, and capturing it within a database for broader access and use.

Modes of knowledge creation

As indicated above, knowledge is created through the interaction between tacit and explicit knowledge. Four different modes that serve to expand knowledge creation are recognized (Nonaka and Takeuchi, 1995):

- *Socialization*: Converting tacit knowledge to tacit knowledge. "Watching, then doing."
- *Externalization*: Converting tacit knowledge to explicit knowledge. "Doing, then describing."
- *Combination*: Combining explicit knowledge with explicit knowledge. "Finding, then combining."
- *Internalization*: Distilling explicit knowledge into tacit knowledge. "Watching, doing, hearing or reading, then believing."

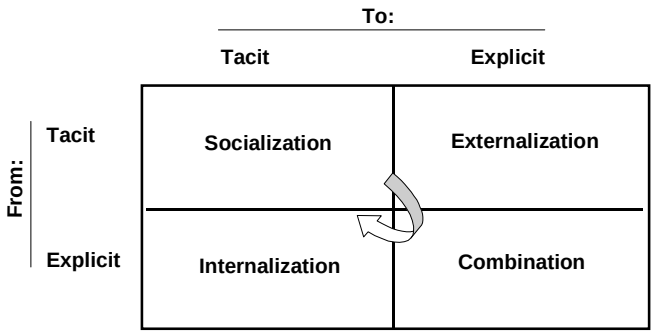
Socialization is associated with group processes and organizational culture. Learning by imitating and doing is how agricultural knowledge tends to be passed from parent to child in rural communities. In the western world, apprenticeships are a form of socialization.

Externalization is a process of converting tacit knowledge into explicit language e.g. by speaking or writing about what one knows. A biological scientist or extensionist working in the field tries to obtain indigenous or local knowledge through this process. Ethnography, individual and group reflection, self-assessment, and systematization (Selener, 1996) exercises also seek to externalize tacit knowledge.

Combination has its roots in information processing. This mode of knowledge expansion involves combining different bodies of explicit knowledge, for example multidisciplinary or participatory approaches to research. Individuals exchange and combine knowledge through such media as documents, meetings, conversations or computerized communication networks.)

Internalization is closely related to organizational learning. This kind of knowledge creation occurs when knowledge and experiences are distilled and internalized, through the combined processes of socialization, externalization, and combination, into individuals' tacit knowledge bases to create shared mental models or technical know-how – very valuable assets to teams and organizations. This is the kind of knowledge creation and communal learning that takes place when stakeholders participate in evaluative activity with experienced evaluators.

Exhibit 6. Four Modes of Knowledge Creation



Source: Nonaka and Takeuchi, 1995.

Ways in which organizational learning takes place

As expressed above, an organization learns through its individual members. Individuals acquire information, make sense of it, and transform it into knowledge. The challenge for the organization is to tap into the knowledge of individuals in order to make use of it to innovate, achieve objectives, and improve performance (Weick, et al., 1993; Weick, 1995).

The term organizational learning, like organizational knowledge, is a metaphor used to capture the idea of knowledge acquisition and use at the organizational level. Organizational learning is engendered within the context of open communication and the shared understandings, experiences, routines, values and behaviors of individuals and groups (Sonnichsen, 2000, p. 76).

Individual learning does not guarantee organizational learning. When people inquire into and reflect upon a problematic situation experienced by their organization they may develop new knowledge, competencies, values, and work routines (Polanyi, 1975; Gephart, Marsick, Van Buren and Spiro, 1996). But for organizational learning to occur, these individuals must be provided with ways of sharing their knowledge in ongoing, systemic ways. To allow this to happen, the organization's managers must make an active commitment to using its members' knowledge and capabilities.

Required are leaders and managers who:

- Value learning and the potential learning brings to the organization
- Promote and facilitate learning in concrete ways
- Provide channels of communication that facilitate frequent and easy access to information
- Establish systems and structures that ensure that learning takes place
- Encourage lessons learned to be put into practice

(Argyris and Schön, 1996; Preskill and Torres, 1999).

Organizational change

Learning is crucial to ensure that organizational change is undertaken to enhance performance. *Organizational change* refers to alteration or variation in the character or performance of an organization. Such changes lie along a continuum from incremental change to fundamental or large-scale change. While incremental change is less complex than fundamental change, both types involve three basic stages referred to unfreezing, moving and freezing (Armenakis and Field, 1993). Fundamental or large-scale change has been defined as lasting change in the character of an organization that significantly alters its performance (Ledford et. al, 1989). Organizational character refers to the design of the organization (strategies, structures, technology configurations, formal information and decision making systems, and human resource systems) and to organizational processes (organizational behavior, communications, decision making, participation, cooperation, conflict, politics, and the flow of information and materials). Performance refers to the execution of activities in relation to objectives.

Organizations are the principal mechanisms through which we organize our environment and fulfill stakeholders' needs. They structure effort and allocate functions to create order out of uncertainty and stability within turbulence. Organizations are confronted with pressures to change, which, if ignored or badly managed, can result in their downfall (Nicholson, 1993). Change has become a way of life for those organizations that are able to thrive in today's competitive environment. Boards of Trustees and managers around the world are concerned about change and about their own and their organization's ability to respond to the increasingly complex demands of an increasingly complex world (Smillie and Hailey, 2001).

Using organizational learning for change and performance improvement

Organizational learning is a core value and a prerequisite to successful change efforts. Research and development activities are essentially knowledge-based and knowledge-sharing processes (Jackson and Kassam, 1997). Knowing what works and why – as well as what does not and why not -- is essential to the success of organizations working in research and development (Smillie and Haily, 2001). The report 'Knowledge for Development' (World Bank, 1998) emphasizes the importance of acquiring, absorbing, communicating and acting on knowledge, if the gap between developed and developing countries is to be narrowed (Smillie and Haily, 2001).

Many organizations fail to learn due to their limited capacity to see past error as a potential source of learning (Fowler, 1997). An apparent inability to learn and a lack of organizational memory are widespread failings of many development organizations. This is partly due to the fact that there are powerful reasons to conceal negative development findings instead of using them as lessons. The increasingly competitive aid environment, the public's waning faith in how well their taxes are used and the increasingly rigid conditions imposed by donors, means that development organizations are discouraged from publicizing – even acknowledging – their mistakes. Thus the potential for shared learning from failure within the development community is reduced to a point close to zero.

Experiences with organizational change in public sector organizations and large corporations (Fullan, 1991; Harvard Business School, 1998; Hobbs, 1999; Hoy and Miskel, 1996; Huberman and Miles, 1984; Mohrman et al., 1989) highlight the importance of the following factors that promote change:

- An external environment that encourages change (e.g., strong external pressures for change).
- Top managers provide leadership for change.
- The change process is adequately managed.

- A critical mass of staff members are involved in the change process and are committed to it.
- Appropriate institutional innovations are provided are developed.
- Resources are provided for change (e.g., dedicated time of key staff members and budgets for training and facilitation).

Before real change can take place, both individuals and organization must "learn how to learn" (Harrison, 1995). This involves continuous reflection *for* action, reflection *in* action, and reflection *on* action (Coats, 2000). The importance of higher order or double-loop learning is emphasized, the complex process which questions basic assumptions and commonly held mental models, and extends the focus of enquiry outside the current set of preconceptions. The process of learning how to learn must be initiated and unwaveringly supported and modeled by management (and led by a dedicated facilitator if managers lack the skill) in order to engage the willingness and ability of employees to involve themselves in collective learning with the intention of accomplishing tangible change targets (Schiemann, 1993).

Organizational performance can be viewed as a function of management's capacity to create a fit between the organization's internal arrangements and external forces (Lawrence and Lorsch, 1967). Many managers have neither the skills nor the training to do much more than keep things going from day to day and so they neglect the potential for mobilizing organizational learning in the pursuit of institutional change and performance enhancement.

The effectiveness of an organization's strategy depends upon two principles: contingency and consistency (Child, 1984). Contingency represents the degree to which the leadership and strategic orientation of an organization are able to match external demands. Consistency represents the degree of coherence with which institutional arrangements, organizational structures, sub-systems and culture are integrated and collectively fit and serve organizational goals (Nicholson, 1993).

Successful change begins with developing a clear strategic direction. Broad participation, consensus building and commitment are pivotal for strategy implementation (Pfeffer, 1993). Active, conscious, intentional organizational change only begins when the strategy 'template' is placed over an organization's structure, systems, culture, and capabilities. At this point, gaps and discrepancies become apparent and a plan of action can be collectively developed for improvement. This process involves

- Strategy development
- Alignment of structures, systems, culture, and capability with the strategy
- Gap analysis
- Action planning
- Implementation

Such a process creates awareness of the need for, and the focus of, significant changes that need to occur. Organizations that have not learned to learn are unprepared for this process, do not possess the required internal culture and ethos for it to thrive and so frequently fail before the process gets underway.

Use of evaluation for organizational learning and change

Evaluative inquiry within organizations is being increasingly adopted by managers and evaluators to help diagnose areas in which change is needed and to help direct the change process (Preskill and Torres, 1999; Lusthaus et al., 1999; Patton, 1999).

Uses of evaluation

It is often assumed that the main use of an evaluation occurs when the conclusions and recommendations presented in the evaluation report are put to use by a manager or policy maker. However, a growing body of knowledge indicates that such *direct or instrumental use* is the exception rather than the rule. In addition to direct use, Weiss (1998) and others have identified two other basic types of use of evaluation findings: indirect and symbolic uses. *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 carried out as symbols of compliance with administrative directives or to present an image of "modernity." Research on evaluation use indicates that symbolic and indirect uses are much more common than the direct use of evaluation findings in decision-making. Direct use is especially rare where politically sensitive issues – such as funding or continuation of a program – are at stake.

Research on evaluation use and organizational learning carried out by Balthasar and Rieder (2000) suggests that the main direct use of evaluation findings is at the operational level of projects and programs. Uses of findings at higher organizational levels are generally indirect.

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 higher decision-making levels is a much more complex process, which requires keeping in touch with decision makers and exploiting "windows of opportunity" when they occur. Such opportunities may occur during an evaluation or long after it has been completed (Balthasar and Rieder, 2000: 257-258).

Evaluators generally think in terms of enhancing the use of the evaluation *findings presented* in their reports. But recent studies and experience indicates that the evaluation *process* can also have an influence on attitudes, 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 subsequent decisions and actions.

Based on a comprehensive review of the literature on evaluation use and on studies and experience with evaluation processes within specific organizations, Forss, Rebien and Carlsson (2002: 33 – 37) have identified five distinct types of process use:

- *Learning to learn*. Learning a structured way of thinking about reality and generating knowledge
- *Developing professional networks*. Being part of an evaluation allows participants to meet with others whom they otherwise have little interaction with.
- *Creating shared understandings*. The usefulness of an evaluation hinges directly upon the quality of communication, which is enhanced through participation in the evaluation exercise.
- *Strengthening the project or program*. Involvement in an evaluation helps participants clarify their understanding of the project or program under scrutiny and may enhance their resolve to achieve its aims.
- *Boosting morale*. Direct and substantial involvement in an evaluation, which brings insights about the strengths and weaknesses of projects appears to boost the morale of participants.

Our own work with ISNAR projects to evaluate organizational capacity development in many countries has provided similar findings.

Barriers to the use of evaluations for organizational learning

There is increasing impatience with development organizations that seem exceedingly reticent or slow to learn lessons from the work they have carried out, the studies conducted, and the evaluations completed. This has led to claims that development studies are irrelevant (Edwards, 1989) and that development organizations suffer from "learning disabilities".

In 1999, the Expert Group on Development Issues in Stockholm convened a meeting in which seasoned development managers and evaluation practitioners came together to discuss "Learning in Development Cooperation." Those present identified three main groups of factors that explain the poor response of aid organizations to failure (Berg, 2000):

The complex and dynamic nature of the development environment. The environment in which development programs are designed and implemented is extremely complex and dynamic, especially where capacity-building is concerned. The technical problems to be tackled are complex. The social and political obstacles are formidable. The conceptions of what is right and wrong are hazy and often controversial. Bargaining and compromise solutions are the rule. Objectives, opportunities and challenges vary from place to place and change rapidly with time. All of these factors make it difficult to draw lessons from experience that might be usefully applied in the future and under different circumstances.

Shortcomings in evaluation. One problem is the sheer number of evaluations. Many thousands of evaluations of development projects and programs have been done in the last two decades, and there has been little attempt to synthesize the main findings of them. Paradoxically, despite the large number of evaluations done, the evaluation reports produced are read by few people. Many evaluations are regarded as confidential. Many others are written in languages or employ terminology that cannot be understood by those evaluated or by those who might be interested in the evaluation. Few evaluation reports are widely disseminated to potentially interested parties. Fewer yet are read by those who receive them. And extremely few are subjected to serious review and debate. Another shortcoming is that many evaluations attempt to meet the needs of distinct groups. They attempt to guide decision-making in aid agencies; to establish accountability; to draw lessons for future programming; and provide a framework for dialogue, including with beneficiaries. The multiple agendas of evaluations, in relation to the resources provided, have frequently led to dissatisfaction with the results. All these factors have conspired to make evaluation reports rather poor means of communicating lessons from experience to potentially interested parties.

Internal organizational factors. The third set of factors concerns shortcomings in the organizations that commission evaluation studies. One problem concerns the rapid turnover of managers and staff members in development organizations; the departure of experienced staff members often means that the lessons drawn from evaluations go out the door with them. Another problem is that information flows – both vertical and horizontal – are sluggish in many organizations, causing evaluation results to remain in the heads of those who conducted the evaluation or those few who read the report. A third factor concerns the "path dependency" of organizations. Individuals and organizations that have strong competence in a particular line of activity will not wish to move away from it, regardless of evidence of past ineffectiveness or the emergence of new demands. Another factor relates to the concern in most development agencies for *technical / analytical questions* as opposed to matters of *process*. Institutional and process issues have generally received very little attention in the planning of development projects. There has been a tendency to overestimate local commitment and capacity and to design overly complex projects. Even though evaluations point out the need to simplify project designs, planners tend to ignore this "process advice" and produce more and more technically sophisticated designs. Due to the political dimensions of leadership in many organizations, there is also a tendency to promote new initiatives at the top, despite previous unsatisfactory experiences with such initiatives with similar initiatives in the same or other organizations. Middle-level managers generally have strong incentives to

"move money" and to apply "standard recipes," irrespective of the specific needs of different countries and cases.

An important pattern that emerged from this meeting was that in development cooperation, most learning occurs among the local "beneficiaries" through engagement in field activities (Suzuki, 2000). Field staff members of development agencies also have opportunities to learn from similar engagements, and often do learn if they are able to suppress their role as expert, trainer or advisor. Those who learn least are usually those in the head offices of the development agencies and in particular the senior managers. Top-level managers learn least of all for two basic reasons. First, they are the least exposed to field activities where the shortcomings of development strategies are most evident and most is to be learned. Second, the risks of mistakes – and of admitting mistakes – inherent in learning are greatest for senior managers and leaders. When development practitioners make mistakes in the field, these can often be safely considered to be opportunities to learn and improve in future programs. But when a top manager makes a mistake, or is seen to have made one, it can result in his or her dismissal. For this reason, due to the potential negative impact of mistakes they might make, organizational leaders tend to be highly risk averse.

Several authors and practitioners (e.g. Preskill and Torres, 1999; Sonnichsen, 2000) offer practical approaches for using evaluation to promote organizational learning and change. These are described briefly in the next section. Engel and Carlsson (2002) go beyond organizational learning and discuss the broader uses of evaluation for societal learning.

Use of evaluation for organizational learning and change

Evaluative Inquiry

Evaluative inquiry for organizational learning and change is grounded in a constructivist theory of learning (Preskill and Torres, 1999). In practice, individuals share their knowledge, experiences, values, beliefs and assumptions, and shape new, shared knowledge which acts as a platform from which action on important organizational issues can be taken.

There are three phases to evaluative inquiry:

1. A focusing phase in which participants determine what organizational issues and concerns the effort will address, who the stakeholders are, and what the guiding questions will be.
2. An executing phase during which participants design and implement the inquiry. They collect, analyze, and interpret data that directly address the guiding questions. On the basis of the interpretation they develop recommendations and then communicate the processes and findings of the report to the members of the organization.
3. An application phase in which participants formulate and select from alternative courses of action, develop action plans. Finally, they implement and monitor the progress of their plans.

During each of these inquiry phases, participants, with the assistance of one or more facilitators, work together and learn through four types of activities: (i) dialogue, (ii) reflection, (iii) asking questions, and (iv) identifying and clarifying values, beliefs, assumptions, and knowledge.

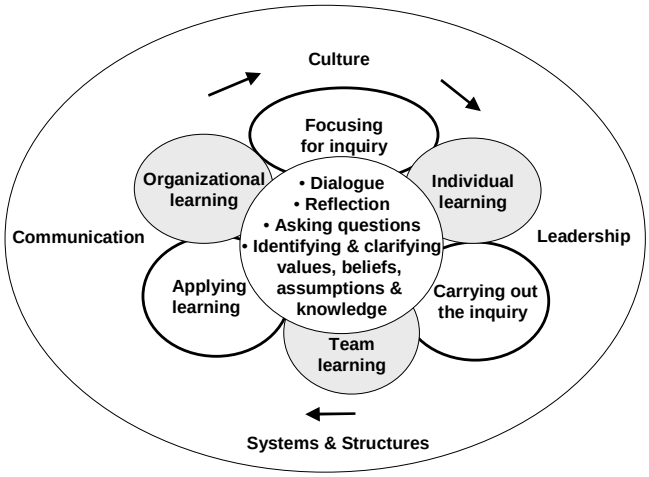
The characteristics of evaluative inquiry include:

- Focus on organizational processes as well as outcomes
- Shared individual, team, and organizational learning
- Coaching of organizational practitioners in inquiry skills
- Modeling the behaviors of collaboration, cooperation, and participation
- Establishing linkages between learning and performance

- Searching for ways to identify and better understand the complex variables that contribute to or detract from organizational success
- Using a variety of perspectives to develop understanding about organizational issues.

In the interests of helping any organization determine its level of readiness for implementing organizational learning, Preskill and Torres have developed the "ROLE" (Readiness for Organizational Learning and Evaluation) Instrument[3]. This short, simple to use questionnaire reflects the findings of current research on how organizations can use facilitative evaluative practices to promote learning and change in the pursuit of excellence. The instrument contains 78 items grouped into six categories that represent critical dimensions known to influence organizational performance: (i) Culture, (ii) Leadership, (iii) Systems and Structures, (iv) Communication, (v) Teamwork, and (vi) Evaluation capacity. The authors report that ROLE is being successfully employed in both private and public sector organizations that include education, health care, manufacturing and high technology. It can be used in its original form or adapted by users to suit their organizational environment (Russ-Eft and Preskill, 2001: 428).

Exhibit 7. Evaluative Inquiry Learning Processes



Source: Preskill and Torres, 1999.

High Impact Internal Assessment

High impact 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 (Sonnichsen, 2000). It uses data-based findings to review the entire range of operational activities engaged in by the organization. It has been effectively used in program improvement efforts and for facilitating and supporting organizational change. It is an action oriented, five-step sequence that focuses on the use of empirical findings to recommend actions designed to improve organizational performance. Sonnichsen's organizational learning sequence framework shows how a five-step evaluative sequence can benefit organizational processes (Exhibit 8).

[3] The ROLE instrument is reproduced in Russ-Eft and Preskill (2001), Appendix A. An electronic version of the instrument can be requested from Preskill at hpreskil@unm.edu without charge, at present.

Following the inquiry phase, evaluation findings are publicized through the advocacy process. This refers to the leaders of learning organizations cogently explaining the reasoning and supporting information that led to their views, actively encourage others to test their views and to provide alternatives supported by the data gathered. Internal communication mechanisms channel the information to organizational members, where it is studied, discussed, refined, and combined with other available information to inform the decision-making. During this stage, the organization is in the learning mode in which its members use new knowledge to inform appropriate courses of action. These are implemented, program direction is altered, and performance modified. Although the flow of events is depicted visually as linear, there is, in practice substantial overlap and feedback among the phases.

Exhibit 8. Five-Step Sequence from Evaluation to Performance Improvement

Source: Sonnichsen, 2000.

4. Implications for Strengthening Organizational Learning and Change in the CGIAR

Strengthening system-level evaluation within the CGIAR

There is evidence that the multi-faceted evaluation system within the CGIAR is being given coherence and, as a result, the dissonance that it has contributed to is likely to be replaced by a greater common understanding and agreement as regards the various kinds of evaluation and the appropriate uses to which each can be put. A proposal for a new approach to evaluation within the CGIAR system has been proposed (TAC Secretariat, 2001). The realization of this proposal has the potential to resolve many of the "evaluation-related" problems described above that currently plague the CGIAR.

In essence, it is proposed that, in the interests of system oversight, the division of roles and responsibilities for all evaluation activities, currently divided between TAC, SPIA and the CGIAR Secretariat, be rationalized. Such a rationalization, the proposal acknowledges, will require several internal developments including:

- A genuine evaluation and impact culture
- An internal (self) evaluation process within centers
- An oversight mechanism, by an independent entity (currently by TAC/SPIA, Secretariat) to ensure the quality and credibility of the internal, center-led (self) evaluation process
- A continued focus upon evaluation for accountability but the encouragement of a parallel process of evaluation to provide information and feedback to centers for learning and improvement
- A safety mechanism to ensure that, while the evaluation system addresses outputs and impacts, that the system's potential for long-term quality of science and research is not compromised or endangered by an over-emphasis on short term outputs and impacts.

These proposals hold the promise to address the unfortunate current tendency within the CGIAR to see the context for center improvement as outside the control of the centers themselves i.e. as residing in the ***external*** and program management review process (our emphasis). This engenders the idea that much of the responsibility for their own quality assurance, accountability and improvement lies outside the pro-active initiative and control of the centers.

The provision of explicit criteria for EPMRs could also tend to empower centers and reduce their dependency on an external review in which they are largely the passive recipients of a process controlled by outsiders and whose criteria are not disclosed in advance. Unfortunately, the proposal has nothing to say about internal program review (IPR). This is a significant omission in any proposal attempting to build a coherent, all-embracing, system-wide evaluation culture.

Despite the dual internal and external review systems within the CGIAR, the process driving them appears to be largely top-down. This is an observation, not a criticism of TAC or the CGIAR Secretariat. With the proposal for a new, integrated and coherent approach to evaluation within the CGIAR, the opportunity to achieve the right blend between direction and oversight from the center and center-led evaluation and development is given a great opportunity. It is yet to be seen whether the proposal will be developed and, if it is, whether centers will grasp this opportunity to institutionalize organizational learning and change in the pursuit of excellence.

In the remainder of this paper, we present some practical implications of our analysis for evaluators and managers in agricultural research organizations.

Practical implications for evaluators

- Design evaluations that focus on the target audience's questions (as opposed to those that interest the evaluator).
- Select methods that best answer these questions (as opposed to adopting the evaluator's preferred, predetermined approach), and that optimize the relevance of the inquiry to organizational members (as opposed to serving the evaluator's professional interests).
- Employ procedures that maximize the involvement and learning of organizational members and stakeholders.
- Use reporting mechanisms and channels that facilitate understanding and assimilation by the intended audience, and that suggest practical ways of using the findings. Be proactive about communicating rather than assume that recipients will read and understand reports.
- Design and execute evaluations as learning exercises, to help participants acquire the discipline of evaluative thinking and expertise.

- Introduce participants to a wide range of evaluation tools, their strengths and weaknesses, so that they can learn to use the right tool for the right purpose.
- Be clear and unambiguous as to the unit(s) of analysis employed in the inquiry: project, program, and/or organization. Explore the complex ways in which organizational levels are interrelated.
- Be sensitive to the relationships between different organizational levels and units of analysis. Organizations are made up of different but interdependent units; what happens at one level can influence what happens at another.
- Be attentive to differences between goal attainment and mission fulfillment. Projects and programs may accomplish their goals and be relatively effective while their contribution to fulfilling the organizational mission is relatively modest. It cannot be assumed that there is a positive, mutually reinforcing relationship between project effectiveness and organizational effectiveness.
- Devote time and effort to defining and strengthening high-quality lessons learned. There is a tendency for "lessons learned" to be appended as a cursory exercise at the close of the evaluation work. Develop criteria for identifying and refining high-quality lessons learned and devote both time and effort to them.

Implications for senior managers

- Ensure that the internal evaluators have a clear and formal mandate to support organizational learning not just to undertake an evaluation and produce a report.
- Use evaluations to engender leadership development. Establish evaluation as a leadership role and function, so that it does not degenerate into a mere reporting activity whose responsibility resides with the evaluation team. Rather than merely add an evaluation unit to the organization, leaders at all levels, need to learn how to incorporate results-based management into their projects and programs.
- Establish overall responsibility for use of evaluation results and lessons at the level of senior management and board of trustee level.

Authors' Brief Biosketches

Douglas Horton is an evaluation specialist and former Senior Researcher at the International Service for National Agricultural Research. His current research focuses on the evaluation of capacity development initiatives and the use of evaluation for organizational learning and improvement.

Viviana Galleno is a former research analyst at the International Service for National Agricultural Research. Her research focuses on the strategic management and development of human capacities for relevant, effective, and efficient organizational performance.

Ronald Mackay is Professor Emeritus at Concordia University, Montreal, Canada, where from 1976 he taught program design, management and evaluation. He works as a consultant in evaluation and organizational diagnostic studies in the public sector and provides training in these areas to government and non-government organizations.

References

Abercrombie, N., Hill, S. and Turner, B.S. (1994). The Penguin Dictionary of Sociology (3rd Edition). London: Penguin.

Alavi, M. 2000. Systems for managing organizational Knowledge. Alexandria: The Organizational Change Program for the CGIAR Centers.

Alston, J.M., Norton, G.W. and Pardey, P.G. (1995). Science Under Scarcity, Principles and Practice for Agricultural Research Evaluation and Priority Setting. Ithaca: Cornell University Press in cooperation with the International Service for National Agricultural Research.

Argyris, C., and D. A. Schön. 1978. Organizational Learning: A Theory of Action Perspective. Reading: Addison-Wesley.

Argyris, C. 1977. Double-loop Learning in Organizations. Harvard Business Review, Sept.-Oct., Vol. 55-5, p. 115.

Argyris, C. 1976. Single-loop and Double-loop Models in Research on Decision-making. Administrative Science Quarterly, Sept., Vol. 21-3, p. 363.

Armenakis, A., and H. Feild. 1993. The role of schema in organizational change: Change agent and change target perspectives. In Handbook of organizational

behavior, edited by R. Golembiewski. New York, N.Y.: Marcel Dekker, Inc.

Balthasar, A. and Rieder, S. (2000). Learning from Evaluations: Effects of the Evaluation of the Swiss Energy 2000 Programme. Evaluation. 6 (3), 245-260. London: Sage Publications.

Berger, P. and T. Luckmann. 1967. The Social Construction of Reality. Garden City: Doubleday.

Bloom, B. 1956. Taxonomy of Educational Objectives. New York: McKay.

Brown, D. 1998. Evaluating institutional sustainability in development programmes: Beyond dollars and cents. Journal of International Development, 10(1): 55-69.

Burki, S. J. and G. E. Perry. 1998. Beyond the Washington consensus: Institutions matter. Viewpoints, Latin American and Caribbean studies. Washington, D.C.: World Bank.

Carlsson, J. and L. Wohlgemuth (eds.). 2000. Learning in Development Co-operation. Stockholm: Expert Group on Development Issues.

Child, J. 1984. Organization: A Guide to Problems and Practice, 2nd ed. London: Harper and Row.

Coats, M. 2000. Learning How to Learn. Milton Keynes: The Open University.

Cooksy, L. J. 1997. Methodological Review and Synthesis of Existing Ex Post Impact Assessments – Report 1: A Review of Documents Reporting Effects of International Agricultural Research Centres. Washington: Consultative Group on International Agricultural Research.

Cooksy, L. J. 1997. Methodological Review and Synthesis of Existing Ex Post Impact Assessments – Report 2: Analysis of Comprehensive Ex Post Studies of International Agricultural Research Centres. Washington: Consultative Group on International Agricultural Research.

Edwards, M. 1989. The Irrelevance of Development Studies. Third World Quarterly, No. 11 (1) January.

Dale, M. 1994. Learning Organizations. In Managing learning, edited by C. Mabey and P. Iles. London, UK: Routledge.

Davenport, T. H. and L. Prusak. 1997. Working Knowledge. Boston: Harvard Business School Press.

EIARD (European Initiative for Agricultural Research and Development). 2002. EIARD Views on Impact Assessment and Evaluation in Agricultural Research for Development.

Engel, P. G. H. 1997. The Social Organization of Innovation: A Focus on Stakeholder Interaction. Amsterdam: Royal Tropical Institute.

Engel, P. and C. Carlsson, 2002. Enhancing Learning Through Evaluation: Approaches, Dilemmas and Some Possible Ways Forward. Background Papers. Maastricht: ECDPM.

Estrella, M., J. Blauert, D. Campilan, J. Gaventa, J. Gonsalves, I. Guijt, D. Johnson, and R. Ricarfort (eds.). 2000. Learning from Change – Issues and Experiences in Participatory Monitoring and Evaluation. London: Intermediate Technology Publications Ltd. and the International Development Research Centre.

Forss, K., C. C. Rebien, and J. Carlsson. 2002. Process Use of Evaluations – Types of Use that Precede Lessons Learned and Feedback. London, Thousand Oaks and New Delhi: SAGE Publications.

Fowler, A. 1997. Striking A Balance: A Guide to Enhancing the Effectiveness of NGOs in International Development. London: Earthscan.

Fullan, M. G. 1991. The New Meaning of Educational Change. London, UK: Cassell.

Gephart, M. A., V. J. Marsick, M. E., Van Buren, and M. S. Spiro. 1996. Learning Organizations Come Alive. Training and Development, Vol. 50-12, pp. 35-45.

Greene, J. C. and T. A. Abma (Eds). 2001. Responsive Evaluation. New Directions for Program Evaluation No. 92, Winter, 2001. San Francisco: Jossey-Bass.

Greene, J. C. and V. J. Caracelli (Eds). 1997. Advances in Mixed-Method Evaluation. New Directions for Program No. 74, Summer, 1997. San Francisco: Jossey-Bass.

Harrison, R. 1995. The Collected papers of Roger Harrison. London: McGraw-Hill.

Harvard Business School. 1998. Harvard Business Review on Change. Boston: Harvard Business School Press.

Hazell, P. and L. Haddad. 2001. Agricultural Research and Poverty Reduction. Food, Agriculture, and the Environment – Discussion Paper 34. August 2001. Washington: International Food Policy Research Institute.

Hobbs, H.1999. Ten tools for managing change in national agricultural research organizations. Research management guidelines No. 5. The Hague: International Service for National Agricultural Research (ISNAR).

Hoy, W. K., and M. B. Miskel. 1996. Educational administration: Theory, research and practice. New York, NY: McGraw-Hill.

Horton, D. 1998. Disciplinary roots and branches of evaluation: Some lessons from agricultural research. Knowledge and Policy 10(4):31-66.

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

Huberman, A. M., and M. Miles. 1984. Innovation up close: How school improvement works. New York, NY: Plenum press.

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

Jackson, E. T., and Y. Kassam. 1997. Participatory evaluation in development cooperation. Ottawa: Kumarian Press.

Jones, A. M. and C. Hendry. 1992. The Learning Organization: A Review of Literature and Practice. London: HRD Partnership.

Lawrence, P. R. and J. W. Lorsch. 1967. Differentiation and integration in complex organizations. Administrative Science Quarterly, Vol. 12, pp. 1-47.

Ledford Jr., G. R., S. A. Mohrman, A. M. Mohrman Jr., and E. E. Lawler III. 1989. The phenomenon of large-scale organizational change. In Large-scale organizational change, edited by A. M. Mohrman Jr., S. A. Mohrman, G. R. Ledford Jr., T. G. Cummings, and E. E. Lawler III. San Francisco, CA: Jossey-Bass Inc., Publishers.

Leeuwis, C. and R. Pyburn (eds.) 2002. Wheel-barrows Full of Frogs – Social Learning in Rural Resource Management. Assen: Koninklijke Van Gorcum BV.

Lincoln, Y. S. and E. Guba. 1985. Naturalistic Inquiry. Beverly Hills: SAGE.

Lusthaus, C., M.-H. Adrien, G., Anderson, and F. Carden. 1999. Enhancing organizational Performance: A Toolbox for Self-Assessment. International Development Research Centre: Ottawa.

Mabey, C. and P. Iles (eds.). 1994. Managing Learning. Milton Keynes: The Open University.

Mabey, C., R. Thomson, and E. Farmer. 1995. The Imperative of Change. In Managing Development and Change. Milton Keynes: The Open University.

Mackay, R., Debela, S., Smutylo, T., Borges-Andrade, J., and Lusthaus, C. 1998. ISNAR's Achievements, Impacts and Constraints: An assessment of organizational performance and institutional impact. The Hague: International Service for National Agricultural Research.

Mackay, Ronald and Horton, Douglas 2003. Expanding the Practice of Impact Assessment and Evaluation in International Research and Development Organizations. ISNAR Discussion Paper No. 03-1. The Hague: International Service for National Agricultural Research.

Maredia, Mywish; Byerlee, Derek; and Anderson, Jock (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, May 3-5, (2000), Rome.

Maturana, H. and F. Varela. 1984. El Árbol del Conocimiento. Las Bases Biológicas del Entendimiento Humano. Santiago de Chile: Editorial Universitaria.

Morgan, P. And A. Qualman. 1996. Institutional and capacity development, result based management and organizational performance. Paper prepared for the Political and Social Policies Division, Policy Branch, CIDA. Quebec, Canada: Canadian International Development Agency (CIDA).

Mohrman Jr, A. M., S. A. Mohrman, G. E. Ledford Jr., T. G. Cummings, and E. E. Lawler III.1989. Large-scale organizational change. San Francisco, CA: Jossey-Bass Publishers.

Nonaka, I. and H. Takeuchi. 1995. The Knowledge-creating Company. New York: Oxford University Press.

OECD (Organization for Economic Co-operation and Development). 1991. Principles for Evaluation of Development Assistance. Paris: Development Assistance Committee.

Patton, M.Q. (2002) Feminist, Yes, But is it Evaluation? In D. Seigart and S. Brisolara, Feminist Evaluation: Explorations and Experiences. New Directions for Evaluation No. 96, Winter 2002. San Francisco: Jossey-Bass.

Patton, Michael Quinn 1999. Organizational Development and Evaluation. In, Andy Rowe (ed.) The Canadian Journal of Program Evaluation, Special Issue.

Patton, Michael Quinn (1997). Utilization-Focused Evaluation: The New Century Text. Thousand Oaks, CA: Sage Publications.

Pedler, M., J. Burgoyne, and T. Boydell. 1991. The Learning Company. London: McGraw Hill.

Pfeffer, J. 1993. Barriers to the Advance of Organizational Science: Paradigm Development as a Dependent Variable. Academy of Management Review, Vol. 18-4, pp. 599-620.

Picciotto, R. and E. Wiesner (Eds.). 1998. Evaluation and Development: The Institutional Dimension. World Bank Series on Evaluation and Development. New Brunswick, USA: Transaction Publishers.

Picciotto, R. and R. Rist (Eds.) 1995. Evaluating Country Development policies and Programs: New Approaches for a New Agenda. New Directions for Evaluation No. 67. San Francisco: Jossey Bass.

Pingali, P. L. 2001. Milestones in Impact Assessment Research in the CGIAR, 1970-1999. Mexico, D.F.: Standing Panel on Impact Assessment, Technical Advisory Committee of the Consultative Group on International Agricultural Research.

Polanyi, M. 1975. Personal Knowledge. In Polanyi, M. and H. Prosch (eds.), Meaning. Chicago: University of Chicago.

Preskill, H. and R. T. Torres. 1999. Evaluative Inquiry for Learning in Organizations. Thousand Oaks: SAGE Publications, Inc.

Quintas, P., E. J. Brauner. 1999. Managing Knowledge and Innovation. In Managing Knowledge. Milton Keynes: The Open University.

Quintas, P., P. Lefrere, Y. Carlisle, and A. Thomson. 1999. Managing Knowledge in an Organizational Context. In Managing Knowledge. Milton Keynes: The Open University.

Rossi, P. H. 1994. The War between the Quals and the Quants: Is a Lasting Peace Possible? In C. S. Reichardt & S.F. Rallis (eds) The Qualitative-Quantitative Debate: New Perspectives. New Directions for Program Evaluation No. 61, Spring, 1994. pp 23-36. San Francisco: Jossey-Bass.

Russ-Eft, D. and H. Preskill. 2001. Evaluation in Organizations – A Systematic Approach to Enhancing Learning, Performance, and Change. Cambridge: Perseus Publishing.

Sanders, J. R. 1994. The Program Evaluation Standards. Thousand Oaks: SAGE.

Schiemann, W. A. 1993. Organizational Change Starts with a Strategic Focus. The Journal of Business Strategy, Vol. 14-1, pp. 43-48.

Schön, D. A. 1983. The Reflective Practitioner – How Professionals Think in Action. London: Maurice Temple Smith Ltd.

Sechrest, L., M. Stewart, and T. Stickle. 1999. A Synthesis of Findings Concerning CGIAR Case Studies on the Adoption of Technological Innovation. Rome: CGIAR Impact Assessment and Evaluation Group, 109 p.

Sechrest, L. 1993. Program Evaluation: A Pluralistic Enterprise. New Directions for Program Evaluation No. 60, Winter, 1993. Jossey-Bass: San Francisco.

Selener, D., C. Purdy, and G. Zapata. 1996. A Participatory Systematization Workbook – Documenting, Evaluating and Learning from our Development Projects. Quito: International Institute of Rural Reconstruction.

Senge, P., M. Roberts, C. Ross, R. B. Smith, and B. J. Kleiner. 1994. The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organization. New York: Doubleday.

Senge, P. M. 1990. The Fifth Discipline – The Art and Practice of the Learning Organization. New York: Doubleday.

Skyrme, D. 1999. Knowledge Management in Practice. In Managing Knowledge. Milton Keynes: The Open University.

Skyrme, D. J. and D. M. Amidon. 1997. Creating the Knowledge-Based Business. Wimbledon: Business Intelligence Ltd.

Smillie, I. and J. Hailey. 2001. Managing for Change – Leadership, Strategy and Management in Asian NGOs. London: Earthscan Publications Ltd.

Sonnichsen, R. C. 2000. High Impact Internal Evaluation. Thousand Oaks: SAGE Publications.

Spink, L. and K. Alison. 2000. Knowledge Management Workshop Beyond Theory: Experience and Opportunities of the CGIAR – Final Report of the OCP Knowledge Management Workshop. Alexandria: The Organizational Change Program for the CGIAR Centers

Stufflebeam, Daniel L. (2001). Evaluation Models. New Directions for Program Evaluation No. 89, Spring, 2001. Jossey-Bass: San Francisco.

Suzuki, N. 2000. What Prevents Development Organizations from Learning? The Difficulties in Learning to Be Learners. In Carsson J. and L. Wohlgemuth (Eds.) 2000, Learning in Development Co-operation. Stockholm: Expert Group on Development Issues.

TAC (Technical Advisory Committee) Secretariat (1997). Terms of Reference and Guidelines for EPMRs of CGIAR Centers. Rome: FAO.

TAC Secretariat (2001). Proposal for a New Approach to Evaluation in the CGIAR System. Agenda Item 6: The CGIAR Technical Advisory Committee, Eightieth Meeting, ICARDA, Aleppo, Syria, 26-30 March 2001. (SDR/TAC:IAR/01/04)

TAC and the Secretariat of the Consultative Group on International Agricultural Research. 2001. The Future Impact Assessment in the CGIAR – Needs, constrains and options. Proceedings of a workshop organized by the Standing Panel on Impact Assessment of the Technical Advisory Committee. Rome: Food and Agriculture Organization of the United Nations.

Uphoff, N. 1986. Local institutional development: An analytical sourcebook with cases. Connecticut: Kumarian Press.

Van der Veen. 2000. Learning Natural Resource Management. In Deepening the Basis of Rural Resource Management – Proceedings of a Workshop. The Hague: RIMISP and ISNAR with the support of IIED, ISG, CIRAD-TERA, INTA, ECOFORÇA, EU and IDRC.

Weick, K. E. 1995. Sensemaking in Organizations. Thousand Oaks: Sage.

Weick, K. E. and K. H. Roberts. 1993. Collective Mind in Organizations: Heedful Interrelating on Flight Decks. Administrative Science Quarterly, Vol. 38, No. 3, Sept., p. 357.

Weiss, C. 1998. Evaluation – Methods for Studying Programs and Policies. Upper Saddle River: Prentice Hall.

World Bank. 1998. Knowledge for Development – world Development Report. New York: Oxford University Press.

Worthen, Blaine, R., Sanders, James, R. and Fitzpatrick, Jody L. (1997). Program Evaluation: Alternative Approaches and Practical Guidelines. New York: Longman.