The role of evaluation in successful integrated natural resource management

Steve Twomlow[1] and Nina Lilja[2]

ICRISAT, P.O. Box 776, Bulawayo, Zimbabwe s.twomlow@cgiar.org
CGIAR System wide Program on Participatory Research and Gender Analysis, CIAT, AA 6713, Cali, Colombia
n.lilja@cgiar.org

Abstract
Evaluating the impacts of NRM research is essential to understand its contribution to increasing agricultural productivity and sustainability, reducing vulnerability, and ultimately alleviating poverty. A proper evaluation helps judge merit or worth, improve programs and generate knowledge. One indicator of a constructive evaluation is that it promotes learning and that the results are used to guide change. This paper suggests that a key to successful evaluation in NRM is in the mindset of the researcher. We discuss the contribution of four factors to a reflective learning process that is necessary in integrated NRM. They are: stakeholder participation, systems approach to evaluation, timing of the evaluation and an iterative approach to investigation.

Media Summary
Evaluation of Integrated Natural Resource Management contributes to learning-oriented approach to enhancing agricultural productivity, sustainability, and alleviating poverty.

Key Words
Evaluation, Integrated Natural Resource Management

Introduction
The significant problems we face cannot be solved at the same level of thinking we were at when we created them. - Albert Einstein (1879-1955)

Evaluation is a broad concept that refers to a systematic assessment of a situation at a given point in time, past, present or future.[1] In the latter case evaluation relies on projections and assumptions rather than on actual measurement or observations. In public sector agricultural research, evaluations are considered an integral part of good management, all too often carried out by external evaluators. Their main goal is to draw lessons from past experience and to incorporate them systematically into the planning and implementation of future activities. At the institutional and systems level, evaluation results enable planners to select or design more appropriate programs and allocate scarce resources among competing research agendas. At the project and individual level, evaluation enables research teams to learn and make project adjustments during implementation, and revise objectives and methods so that the project makes a real contribution to development. In practice, many persons involved in conducting the evaluations are skeptical about whether such evaluations really promote individual, project or institutional learning. The way most evaluations are commissioned and conducted still aims at making definitive judgments about project worth rather than providing learning opportunities. As our approach to solving agricultural and environmental problems changes, so should our approach to evaluation change!

To date much of the work in agricultural research and development has focused at the plot and farm level, in search of the "silver bullet" to alleviate poverty. Public sector research has focused on specific grain and root crops and studied food production processes/supply at a national and international level, rather than on rural food producers and consumers at a local level. As a consequence, there has been progress on short-run issues of improving food production (e.g., the Green Revolution), but we have failed to

[1] The terms impact assessment; monitoring and evaluation are often used interchangeably. A classical way to make a distinction between the three is based on timing, analytical level and specificity. Monitoring occurs frequently, evaluation periodically, and impact assessment usually at the beginning or the end of the project. Monitoring is mainly descriptive, recording inputs, outputs and activities. Evaluation examines processes, while impact assessment is concerned with long-term outcomes. Monitoring is specific and compares planned and achieved results, evaluation does the same but also looks at processes. Impact assessment is less specific and also considers external influences and events.

address long run social and ecological needs of smallholders in more marginal environments (Twomlow, 2003; Ashby 2001).

Ian Johnson, chairman of the Consultative Group on International Agricultural Research (CGIAR), has observed that mismanagement of natural resources is the "Achilles heel" of long-term sustainable development (CGIAR, 2000). One of the major lessons learned is that the lack of participation by beneficiaries of a project, at the design stage, contributes to project failure. Suggesting that researchers, extension and development communities work together and be aware of inter-sectoral linkages is nothing new. What has been missing is an effective framework that allows research to better accommodate the full range of factors and players and be aware of the nature, causes and potential results of conflicts and constraints within agro-ecosystems. Integrated natural resource management (INRM[2]) is an attempt to build a new agricultural research and development paradigm to meet this challenge. Campbell et al. (2001) define INRM as "a conscious process of incorporating the multiple aspects of natural resource use (be they bio-physical, socio-political or economic) into a system of sustainable management to meet the production goals of farmers and other direct users (e.g. food security, profitability, risk aversion) as well as the goals of the wider community (e.g. poverty alleviation, welfare of future generations, environmental conservation)."

Researchers are increasingly using participatory methods and hence many projects are designed to be responsive to changing community needs. Recent research on the poverty alleviating impacts of technology associated with the CGIAR has identified institutional learning and change (ILAC) as a key area for intervention if research is to be more effective in serving the poor (Hall et al., 2003). ILAC is a process that can change behavior and improve performance by reflecting on and reframing the lessons learned during the research (Horton and Mackay, 2003). Sayer and Campbell (2003) argue that monitoring and evaluation is the key to the adaptive project management and reflective learning required for successful INRM, but few know how to respond in ways that generate relevant information for those involved in INRM type initiatives.

This paper will not discuss the principles and practices of INRM. Nor will it attempt to provide a roadmap to successful evaluation practice because there are number of published documents already available (Barrett et al., 2003; Campbell and Sayer, 2003; CIMMYT, 2003; Douthwaite et al., 2003; Harwood and Kassam, 2003; Perez and Tschinkel, 2003; Pound et al., 2003; Shiferaw and Freeman, 2003; Agricultural Systems Vol: 78). Several good manuals and guides exist on how to design and plan an effective evaluation process (e.g. Guijt, 1998; Herweg and Steiner, 2002; OECD, 2002). The purpose of this paper is to discuss what we call the "mindset of the researcher" which relates to issues that are relevant and unique to the evaluation of INRM research, and to discuss how paying closer attention to these issues and adapting scientific methods contributes to a reflective learning process that is key to successful INRM. This paper will discuss four issues in the evaluation of INRM research: stakeholder participation, systems approach to evaluation, timing of the evaluation, and iterative approach to investigation.

The NRM Challenge

INRM grew out of farming systems research (FSR), which had its heyday in the mid 1980s and then all but disappeared from the list of research programs by the early 1990s. FSR attempted, just as INRM is attempting today, to carry out research with complicated technologies in complex settings. Research on complex agricultural systems is difficult because of the multiple scales of interaction and response within and between physical and social subsystems, uncertainty, long time lags, and multiple stakeholders with often contrasting objectives and activities (Campbell et al., 2001). Early FSR failed because by engaging with this complexity it was criticized for generating excessive amounts of data, for being very costly to conduct and yielding few results of immediate practical value or impact. The other major cause of the failure of FSR was a lack of understanding of the role and power relations of farmers and other stakeholders in technology development (McCown, 2001). In many cases, researchers conducted their experiments in farmers' fields but without sufficient interaction with the farmers themselves. In other words, they continued their traditional research methods but only outside the experimental station, and the way research was organized was never questioned. Participation of private firms, consumers and farmer associations in the planning and execution of research was minimal. Research systems and institutions

[2] INRM is assumed to include all efforts in integrated genetic resource management.

were rigid, and unable to change the way research was organized and conducted. Early FSR learnt from its mistakes, and NRM research evolved to pay more emphasis to stakeholder participation. But many

Box 1: Common organizational problems in NRM research (adapted from Ashby, 2003)

- Key stakeholders are not represented in the research process
- Participation is not developed around clearly specified rights, roles and responsibilities
- Mechanisms of accountability among participants (especially among researchers) are lacking,
- Process too often corrupted by hidden agendas
- Conflicts of interest are not made explicit or negotiated
- Transaction costs of participation exceed the benefits to the participants
- Feedback mechanisms, such as monitoring and evaluation of the research process, are not in place. Consequently, learning about how to improve the process is minimal or slow.

organizational problems still existed (Box 1). Eventually INRM resulted of this process of trying to improve the effectiveness of NRM research.

A recent survey of 59 participatory NRM projects indicates that today many projects tend to take an integrated approach to NRM, looking at multiple resources and developing biophysical and institutional innovations for improving their management. Over half the projects in the survey worked at multiple scales or at scales defined by social rather than biophysical criteria alone (Johnson et al. 2000).

In essence INRM tries to harmonize the complementary but often conflicting goals of production and environmental protection. The current thinking on INRM can be described through a conceptual and an operational framework (Figure 1). Within the conceptual framework there are three pillars:

- Addressing issues of what type of science to do where
- Changing the social organization of science
- Establishing a system for adapting and learning.

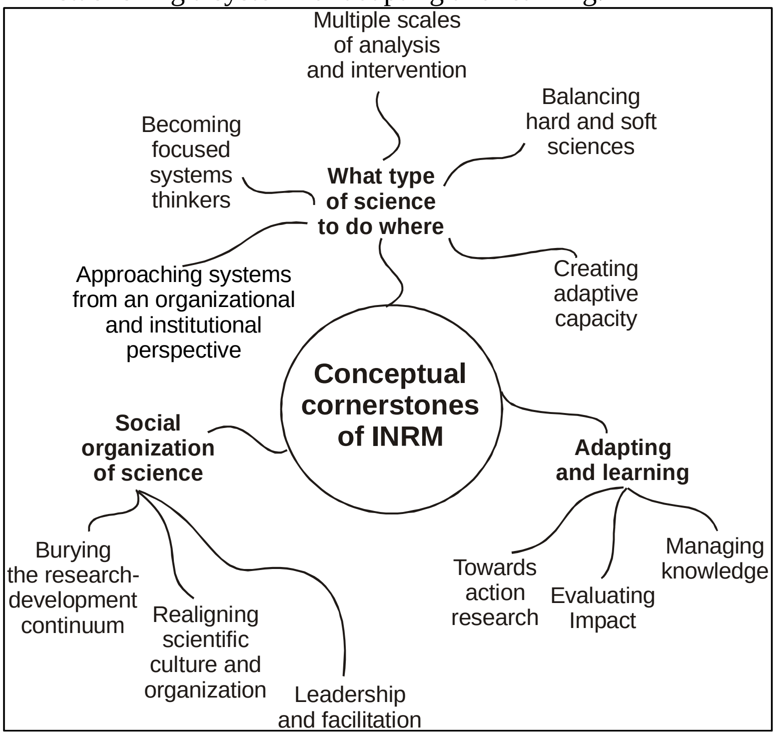

Figure 1. The conceptual pillars of INRM (Campbell et al., 2003)

It is argued that, to bring science to bear on poverty and sustainability simultaneously, society needs a new social contract for science. For putting INRM into action, an operational framework consisting of 11 ‘cornerstones’, which can be used as a guideline for the implementation of INRM projects, has been developed. The cornerstones can be used as a checklist for self-reflection and evaluation, such that each cornerstone needs to be considered; otherwise the weakest becomes a threat to the whole. One of the biggest challenges is not becoming lost in the complexity.

The several unique features of NRM technology create methodological difficulties for NRM project evaluation. Unlike germplasm technologies, the impact of NRM technology occurs only indirectly through the economic and environmental goods and services that generate direct and indirect benefits to society. These benefits are often multi-faceted, including economic, environmental and social gains across space and time. Hence, these benefits are often externalised, and may not be entirely captured by the investor, or even documented in a ex-post impact assessment, until many years after the work was completed (Box 2).

Box 2: Attributing impacts on grain yields in South America (adapted from Ekboir, 2003)
In the 40 years between 1961 and 2001 production of maize, sorghum, sunflower, soybeans and wheat in MERCOSUR increased from 23 million to 152 million tons. The increase came about because farmers adopted three interdependent technologies: the introduction of soybeans in the late 1960s, zero tillage, and improved germplasm. Soybean production led to an intensification of agriculture, which caused serious soil degradation.

A number of technical solutions were proposed to combat degradation, including zero tillage and terracing. At the time, researchers identified terracing as the more promising option, and as a result soil conservation projects neglected work on zero tillage. This was partly because the research on zero tillage, which started in the early 1970s, used the herbicide Paraquat that was difficult to use. The zero tillage package only became technically feasible after the release of glyphosate (Roundup) in 1976. No ex ante impact assessment could have predicted the existence of Roundup before 1976, and if research managers had assumed that something similar would be invented, it would have been dismissed as fanciful. Adoption, however, remained low until the early 1990s because glyphosate was expensive. Then, a change in corporate policies helped reduce the price from US$ 40 to less than US$ 10 per litre. This, together with a very effective diffusion campaign organized by farmers' associations, caused adoption to explode. The impact of these technologies cannot be separated. Without zero tillage, the impact of improved germplasm would have been very small, as zero tillage was necessary to stop soil erosion and improve water management. At the same time, new and improved germplasm increased the profitability of zero tillage, fostering adoption. But adoption only became feasible with the development of glyphosate and only really took off when it became substantially cheaper.

It is becoming increasingly clear to those evaluating the full impact of INRM research and development programs, that they must add appropriate indicators of both social and natural resource endowments and well-being to the limited, traditional economic indicators if they are to accurately assess impact (Harwood and Kassam 2003, eds.). Kelley and Gregersen (2003) paint an even bleaker picture about some of the methods needed: "When addressing NRM research impacts, a whole range of other issues needs to be considered. Markets are largely missing for the environmental services provided. Different valuation methods exist, all of which are highly imperfect and tricky to use, and hence need bracketing attributing prices from different angles. Externalities are spread over different scales and hence difficult to capture as each level needs to be done with different tools. The time dimension is crucial and hence the choice of discounting is key. There are also important problems of resilience and irreversibilities that need to be taken into account in constructing counterfactual scenarios. For these reasons, designing control groups for NRM treatments is particularly difficult because of the spatial and temporal dimensions involved". NRM is like jazz, or raising children: It requires constant improvization and there is no single correct way of doing it. NR managers are constantly confronted with surprises. Stakeholders change their aspirations and exogenous factors have unpredictable influences on the system (see Box 2).

Four issues need to be addressed:
- Stakeholder participation
- A systems approach to evaluation
- Timing of the evaluation
- Iterative approach to investigation.

Issue 1: Stakeholder Participation
Although it is widely agreed that evaluation is a useful and necessary process for project learning, many researchers often take the path of least resistance where evaluation is concerned, and evaluation plans are seldom made at the project design stage. There is a perception that manpower for evaluation is lacking; project team members often believe that only experts in evaluation can do it. Driven by an accountability motive, there is also deep-rooted trust that external evaluations are more objective than those carried out by internal project staff. The latter are sometimes discounted as not objective or even biased (Shiferaw and Freeman, 2003). It is erroneously assumed that evaluation must be a management initiative, and this misconception enforces the notion that stakeholders in a project (other than donors and management) cannot objectively evaluate their efforts. Because of this quest for objective assessment by an outsider, the judgment function of the evaluation dominates its learning function.

There is wide recognition that participatory approaches are critical for achieving sound NRM, and that stakeholder participation is central to the approach. Yet to date most of the available evidence is frequently descriptive with little evidence of impact or benefit to date (Pound et al., 2003; Barrett et al., 2003). It is often assumed that stakeholder participation extends to evaluation activities as well, but that is not necessarily the case. In the emerging era of institutional learning and change, full involvement of the project stakeholders as evaluators, and not only as those involved in evaluation activities, is crucial. It is in fact a necessary condition in the evaluation process, and is meant to lead to learning and change at any level, whether individual, project, institution or system level. Participatory approaches to evaluation, particularly in participatory plant breeding, have effectively highlighted the importance of using stakeholder-defined indicators, and generally more intensively involving a wider range of stakeholders in providing information for the evaluation (e.g. Ceccarelli et al., 2000). In practice the organizational system that supports research does not inherently encourage stakeholders to act as evaluators.

An evaluator needs to keep in mind that all evaluations are done for a specific audience. The evaluation answers specific questions that are of interest to the audience. The questions are answered using data that can take many forms; this is an empirical process that uses scientific methods. Identifying the audience for the evaluation can be difficult, especially in NRM research where there are numerous people (even apart from project stakeholders and beneficiaries) who are interested and need to use evaluation results. Any given project evaluation has multiple stakeholders, that include the project team; the donors; project collaborators and beneficiaries, and the evaluation team as illustrated in Figure 2.

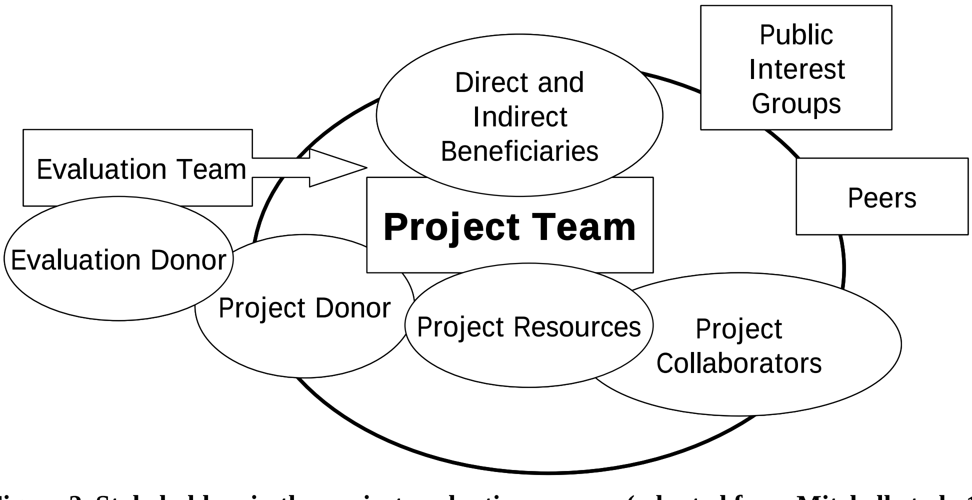

Figure 2 Stakeholders in the project evaluation process (adapted from Mitchell et al., 1997)

In the context of evaluation it is important to consider the dynamics of stakeholders. Stakeholders and their stakes may change over the course of the project. Stakeholders can be characterized according to three attributes they may or may not possess: power, legitimacy, and urgency (Mitchell et al., 1997). A stakeholder salience depends on how many of the three attributes they possess. Obviously different stakeholders have different needs from evaluation, and the evaluator needs to clearly understanding who is the audience of the evaluation, what are dynamics between stakeholders, and most importantly the salience of each stakeholder which may dictate the evaluation process – for example, influencing what type of questions are investigated, which indicators are used, and how success or failure are defined.

Generally we expect too much from evaluation because we expect studies to be definitive, to tell us exactly what to do, or to pass an ultimate judgment on the project success, and for results to appeal to a universal audience. In doing so we set up ourselves to fail. If we view evaluation as a key factor in the process of adaptive resource management and institutional learning and change, then "successful" evaluations need only to be helpful (in terms of how well they support learning and change) to an identified audience for which the study is performed.

Issue 2: A Systems Approach to Evaluation
Regardless of who is the audience, each evaluation needs to be viewed as contributing "actionable" assessments corresponding to the three INRM pillars defined above, i.e. (a) what type of science to do where, (b) the social organization of the science, and (c) adapting and learning system. Thus, the evaluator needs to focus on three systems simultaneously: the science-system, which develops the technologies, the organizational system that supports the research, and the system that combines the two and fosters learning and change.

Currently the most dominant use of evaluation results is in contributing to the type of science we do. We now accept that standardized, widely applicable technologies are unlikely because small-area producers generally have multiple objectives; and because achieving impact involve the interplay of multiple stakeholders, often beyond the duration of a project (see Box 2). Rather, research efforts should be directed at improving the capacity of agroecological systems to adapt to changes and to continue to supply a flow of products and services that poor people depend upon, i.e., to improve the "adaptive capacity" of the target system. In practice this means helping farmers and other managers of natural resources to acquire the skills and technologies to better control their resources, i.e., improving their adaptive management abilities. INRM's way of working is to develop practical, local solutions with farmers together with an array of local and international partners. In deriving the solutions the best science is blended with local and specialized technological knowledge. The underlying principles learned in the local process can then be used to develop solutions for similar conditions elsewhere. Scaling up of project impact means scaling up the process of fostering adaptive management, not extension of universally applicable management practices or technologies.

In an analysis of more than 200 cases of sustainable agriculture from 52 countries, involving more than 9 million farmers farming more than 30 million hectares, Pretty and Hine (2001) concluded that success occurred when participatory approaches were used that involved farmer experimentation and built capacity to learn about biological and ecological complexity. A major challenge in INRM is to combine the various "information bits" derived from different stakeholders, and distil these into decision rules that they can use (Snapp and Heong, 2003). Hence the evaluators must strive for "adding value" to science. In practice this means ongoing interaction, from the onset of the project, dialogue with users about what type of evaluation information would be useful to them. For example, Syrian farmers' selection criteria for barley, largely based on environmental factors, were quite different from those used by the national breeding program, and often more effective in the more marginal environments (Ceccarrelli et al., 2000). In the process of evaluative investigation, whether it is to evaluate questions regarding what type of science was done or should be done, or assessing the effectiveness of the institutional arrangements and power relations in questions, the evaluator needs to acknowledge the difficulty in determining whose opinion really dominates. A major question, when one moves along the continuum (Figure 3) from conventional or collaborative (Researcher-Led) to more participatory or collegial research approaches (Farmer-Led), is who controls and makes decisions about the research process, and who consequently identifies the milestones and indicators that will be used in evaluation (McDougall and Braun, 2003). This is a real concern because a survey of participatory natural resource management projects indicates that the vast majority of projects report that they do consultative or collaborative research. Relatively few projects report collegial participation at any stage, which shows that while researchers are willing to share control with users, they are not inclined to cede decisions-making authority to them (Johnson et al., 2000). The same study also assessed whether the research reflected a clear and coherent common agenda among stakeholders and if it contributed to partnership building. The study found a lack of consistency in decision-making in terms who makes the decision (researchers or other stakeholders) from one stage of research to the next, which could undermine stakeholders' sense of ownership of the process. Also, the way that participants are selected suggests that the research agenda that emerges may not reflect all points of view, and may be particularly unrepresentative of priorities and concerns of marginalized groups. For example, only 2.1% of the projects selected participants exclusively on the basis of equity criteria.

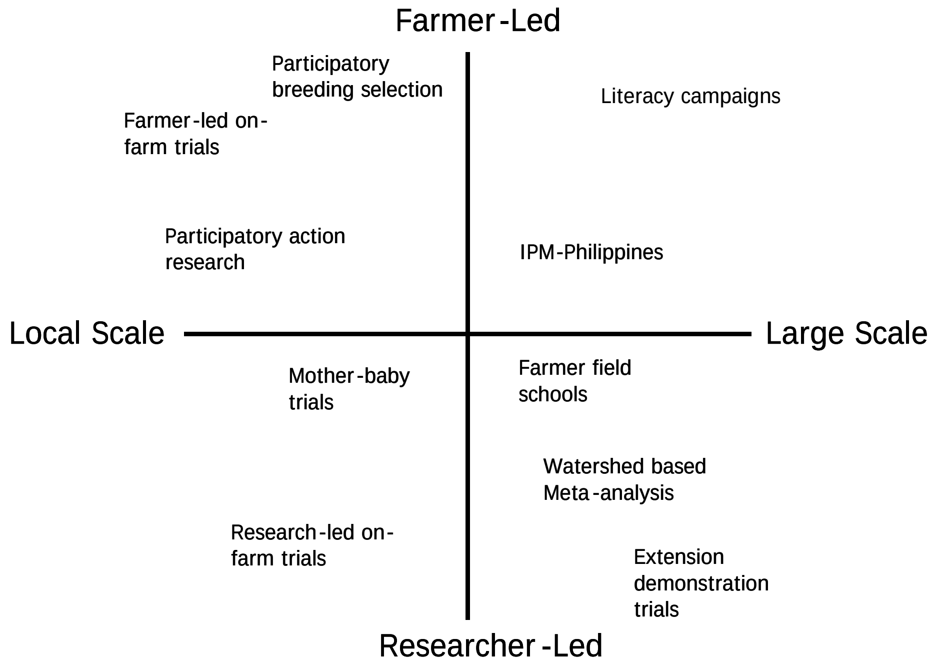

Figure 3. A comparison of participatory learning and research approaches in terms of scale of operation, and degree of farmer versus research involvement (Snapp and Heong, 2003)

All too frequently, as observed by Adesina and Chianu (2002) in an evaluation of alley farming in Nigeria, researchers focus on the biophysical characteristics of the agricultural system and neglect the socioeconomic factors. This leads to inappropriate targeting of the technology and a lower likelihood of adoption and subsequent impact. Early evaluation of the alley farming technologies would have enabled researchers to work more closely with rural communities to adapt the technologies to their specific needs. For example, in many parts of West Africa, women do not have secure land and tree tenure, due to a patrilineal inheritance system. Therefore the technologies were not gender neutral and required modifications for female-headed households. In addition it is important to use spatial analysis to better understand where to target the technologies relative to incentive structures across villages and communities, as determined by market and non-market factors. In fact the most important part of the whole adoption process is the ability of farmers to use their knowledge to modify and adapt technologies. When trying to understand adoption decisions, researchers should make sure they spend enough time evaluating the entire sequence of the adoption process from initial adoption to technology modification and adaptation. A protective researcher/institutional environment can stifle farmer creativity and innovation.

A good example of how a protective institutional research framework can delay the spread of innovation is the case of Zephaniah Maseko Phiri in Zvishavane District in Zimbabwe (Murwira et al., 2001). Phiri developed a number of innovations in soil and water conservation, but it took more than 15 years to spread beyond his farm, as government services providers viewed them as a threat to the country's policy of natural resource management. The technologies disseminated to farmers had to be tested and proven under researcher management, a protective condition that still dominates research and extension in many countries. It was only with the advent of farmer participatory research techniques and a demand for alternative soil water conservation methods, that researchers began to document Phiri's experiences and provide a platform for him to share them with other farmers, scientists and extensionists. Lessons from this work led to the rapid farmer-to-farmer extension of rainwater harvesting (run-on) orchards in a small dam and community resource management project in semi-arid Zimbabwe (Ellis-Jones et al., 2001). In fact the demand for fruit trees, beyond the initial 7000 trees supplied by the project to six pilot communities, required the development and promotion of community based nurseries to meet the demands from more than 100 communities in less than 18 months.

Issue 3: Timing of the Evaluation
Evaluation provides learning opportunities for all stakeholders: country leaders and decision-makers, public sector managers, project manager, team leaders and individual scientist. The key is to look for insights – from both project successes and project failures. These learning opportunities and insights can be derived throughout the project learning cycle (Figure 4). The important notion here is that the learning cycle should not be confused with a project cycle; they are not necessarily the same. Public agricultural research is usually conducted within a given project timeframe. For example, a recent survey of 59 participatory natural resource management projects indicates that the average NRM project length is 4.2 years (Johnson et al., 2000). The "learning cycle" in INRM has important implications to evaluation, as it indicates the timing of the critical input that the evaluator needs to provide. The same study reports that forming of feedback links and changing research priorities were the most common impacts on the research process, with 62 percent of projects reporting them.

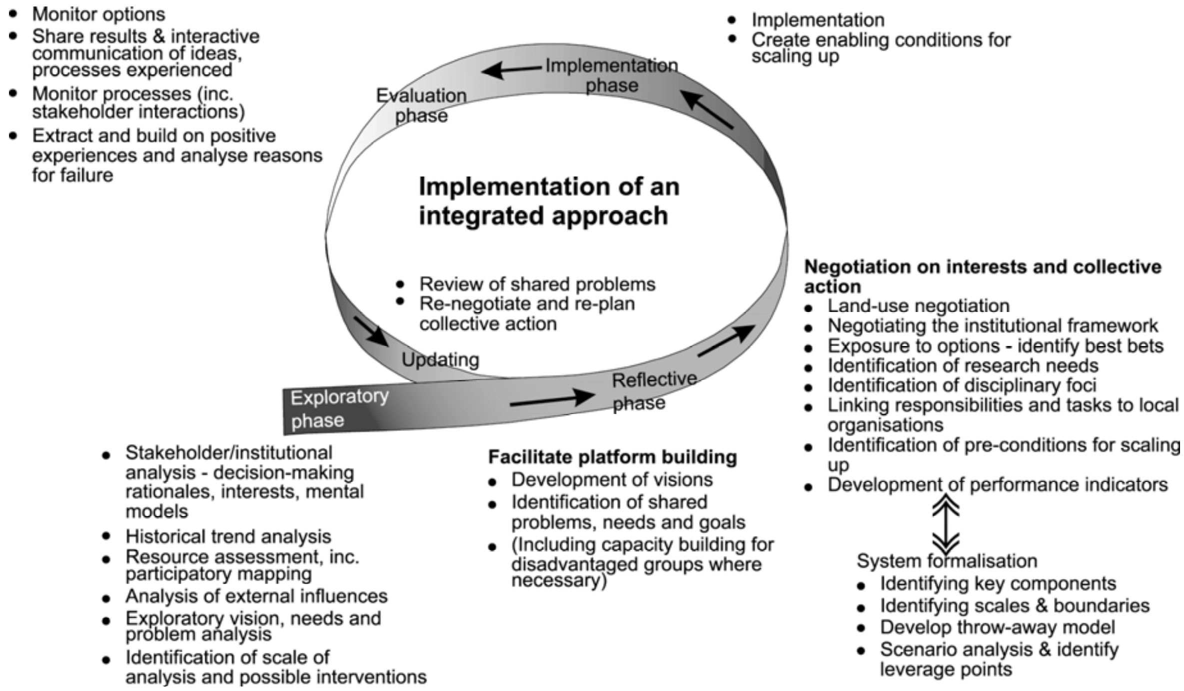

Figure 4. The learning cycle in INRM, where stakeholders undergo reflection-implementation-evaluation. Some of the possible key activities at each of these stages are indicated (Sayer and Campbell, 2003).

Issue 4: Iterative approach to investigation
We can define five stages in the evaluation process. Evaluation begins with negotiation about who is doing the evaluation. Implicitly the decision is also made about the audience, based on who is doing it and who is funding the evaluation. The next stage is defining the evaluation questions. Effective learning-oriented evaluation questions pay attention to both the project task and the process (how it was conducted), and takes a systems approach to evaluation as discussed earlier in this paper. For example, the *task* can be defined as what those involved have to do (e.g. reduce soil erosion). The *process* is concerned with how people and groups work together and maintain relationships. Because task and process are linked, it is important to measure the progress of both.

The investigation stage of the evaluation has traditionally taken a linear approach. This approach assumes that an investigator can measure predetermined indicators of achievement, and that the results should be independent of the observer (no learning will occur for investigator or people involved in the project as they are not involved in the evaluation). A linear evaluation process moves from designing the study, to selection of instruments, data collection and analysis, with very little effort on reflection and re-orientation of the assessment. The linear approach assumes that precise indicators can be defined because everyone agrees or can be brought to consensus on what is good and right. It does not acknowledge that the stakeholders are many and they have different levels of power, legitimacy and urgency.

Reddy and Soussan (2003) attempted to assess the impact of watershed development programmes in the context of the Sustainable Livelihoods Framework. They concluded that assessing the impacts of participatory watershed management using the SRL framework is a methodological challenge as it requires monitoring the changes in the five capital assets, some of which are difficult to quantify in the short term. Hence, the evaluation process needs to be balanced between qualitative and quantitative aspects, as well as long and short-run aspects. A further complication is the one of scale, some indicators are measured at the household level and some at village, community and even national levels. Moreover, attributing change to a particular intervention or programme is difficult, as there could be variables external to the project influencing these changes. For instance, changes in educational and health status could be due to other programs, but they may influence the impacts of the INRM intervention, or vice versa.

Methodological difficulties for NRM impact assessment are rooted in several unique features of such technology interventions. Unlike germplasm technologies, the impact of NRM technology occurs only indirectly through the economic and environmental goods and services that generate direct and indirect benefits to society. These benefits, as donors are beginning to recognise (DFID, 2003) are often multi-faceted, including economic, environmental an social gains across different scales. Hence, these benefits are often externalised, and not entirely captured by the investor.

Evaluation of INRM research calls for an iterative approach that allows room for more stakeholder reflections before data collection, and during and after the analysis. It should be based on the fact that there is no single "truth" about the assessment – different observers (whether individuals or groups) legitimately hold very different perspectives on what is good and right. Iterative investigation also views verbal description (yet not discounting quantitative methods) as essential to portraying these different perspectives. Emphasis on iterative evaluation does not mean less numerical, quantitative assessment, but it allows the data to be put to better use; the data gain value if questions are answered, recommendations are made, and improvements are suggested during the course of evaluation.

Several tools have proved useful in facilitating learning and the innovation processes during NRM research. These include farmer field days and various participatory approaches such as matrix ranking and focus group discussions of technologies under test. These approaches are useful not only for researcher learning but also for farmers to learn from each other. Indeed, one conclusion is that perhaps more emphasis needs to be put on encouraging farmer M&E through field days, exchange visits, focus group discussion etc., as observed during the on-farm testing of a series of 'Best-bet Legume technologies using the Mother-Baby Trials approach in Malawi (Rusike et al., 20041). The results highlighted to researchers the fact that farmers do not make decisions based on agronomic or economic considerations alone (Table 1). For example, researchers initially ranked groundnut-pigeon pea and maize-*Tephrosia* intercrops as the best for farmers because of their high grain yields. Baby-trial farmers, however, ranked maize-pigeon pea intercrop as the best because of the grain-legume mix and the lower labour requirements. According to the baby-trial farmers, the pigeon pea-groundnut rotation was attractive, but only for commercial farmers who had enough land for rotations.

Table 1: Ranking of acceptability of technology options tested in Malawi (adapted from Douthwaite et al. 2003)

Best Bet Option	Agronomic[1] acceptability	Economic[2] acceptability	Farmer [3] acceptability
Unfertilized maize	5	6	5
Maize + area specific fertilizer	2	4	7
Maize+pigeonpea	3	2	2
Maize+pigeonpea+area specific fertilizer	1	3	6
Groundnut+pigeon pea	6	5	3
Maize+*Tephrosia*	4	7	4
Mucuna-maize rotation	7	1	1

1 - Agronomic acceptability ranked in terms of yield performance
2 - Economic acceptability ranked in terms of marginal rates of return analyses
3 - Farmer acceptability based on seasonal matrix ranking exercises

'Expert' opinion has long held that farmers' decision-making is based primarily on economic and agronomic performance. The adoption survey, field days and focus group discussions helped researchers appreciate that farmers' access to resources and climatic risk are equally important criteria. The work helped show the research team that future work in soil fertility management should take into account the poor market for legumes and the fact that different household have different access to resources and hence favor different technology options. The work has also shown that the way field research is carried out, in terms of whether farmers can experiment on their own, the number and types of field days held, and whether training to build the capacity of farmers as experimenters is provided, can greatly affect adoption, adaptation and eventual impact.

Conclusions

Effective evaluation is a requisite part of the succesful project cycle, generating significant impact through project activities and dissemination of results, but it is not on its own a sufficient condition for effective project implementation. Evaluation procedures have been formalized for different types of agricultural research. While they may be partly adequate for some types of research, such as genetic enhancement, the evaluation approaches commonly used are far from adequate – and often entirely inappropriate – for measuring the impact of NRM research. Unlike germplasm technologies, the impact of NRM technology occurs only indirectly. These benefits are often multi-faceted, encompassing economic, environmental and social gains across space and time – usually extending far beyond the project cycle, and therefore “outside the scope” of a conventional evaluation that looks at a specific time frame linked to the project life cycle. Thus, to measure the full impact of INRM research and development programs, evaluators must consider social and natural resource endowments and well-being, in addition to traditional economic indicators. They must also resolve a host of methodological issues: the techniques and tools used, discount rates to be applied, how to apportion impact among different (planned and unplanned) outcomes and factors external to the project.

Apart from highlighting these methodological issues, this paper raises even more fundamental questions of approach. Traditionally, evaluation of NRM projects has focused simply on evaluating adoption or impact. But evaluations can, and should, serve another equally important function, that of encouraging and supporting learning and institutional change. Rather than a traditional “linear approach” to evaluation, where external evaluators measure progress or impact against predetermined milestones or objectives, we suggest a more flexible, more iterative participatory process. Such a process would involve the full range of stakeholders as evaluators. It would involve a conscious effort to reflect, share ideas and perspectives during (not after) the evaluation, and re-orient the evaluation process to better reflect the different perspectives of stakeholders.

In recent assessments of watershed projects in developing countries, Perez and Tschinkel (2003), argue that many of the projects implemented over the last 25 years tried to combine poverty alleviation and resource conservation goals, neither of which have been satisfactorily accomplished. All too frequently the studies were heavily supervised, with prescribed packages and up 100% subsidies awarded to adopters. As such the estimate of adoption is not meaningful, as ex-post analyses carried out on Indian Watersheds suggested that many farmers abandoned the packages once project subsidies ended (Kerr and Chung, 2001). For interventions to have a positive effect it is concluded that a holistic systems approach is required that includes all stakeholders in the INRM process, rather than only the poor farmers, and facilitates a learning process, treating both large and small farmers, as informed clients to whom the research and development organizations are accountable, rather than only to the donors.

Unfortunately, few of the NRM projects today are implementing an evaluation process that takes full advantage of the learning function of the feedback loop (Figure 4), even when it is part of the project activity. A recent survey of 59 participatory NRM projects found that only thirty eight percent reported a change in practices at their research program or system level (Johnson et al., 2000). Clearly, opportunities to learn from the evaluation process, and to consequently change systems of operation, are being lost.

An effective evaluation system for INRM is more than a statistical task or an external obligation. It must be planned, managed, and provided with adequate resources. To ensure that research projects generate impact, the evaluation system must be linked with overall project strategy and operations, and with outputs, outcomes, and impact, and integrated within the project logical framework. However, to achieve this we must change the mindset of the evaluator/researcher to take on board the four issues addressed in this paper:

- Stakeholder participation at all stages of the project and resultant learning cycle is imperative, and it is not the sole responsibility of Project or Institutional management to initiate evaluation activities.
- Systems approach that provides actionable answers to the three pillars of INRM.
- Timing of evaluation – evaluation cycle and project cycle are not necessarily the same thing and different methods will be required for different stages.
- Iterative approach to investigation – formative rather than summative evaluation that may not reach a consensus.

Finally, evaluation methods appropriate for INRM are very different from conventional evaluation practice carried out in research institutions and development organizations. Whether INRM-type evaluation becomes more common will depend largely on donors making scientists and change agents accountable not for impact in unrealistically short time-periods, but accountable for learning, adapting and achieving outcomes that are known to contribute to development. The signs are positive. IDRC (Canada) and GTZ (Germany) and DFID (UK) have started to make the change, not just for INRM but also for all types of integrated development project. The CGIAR Institutional Learning and Change Initiative, supported by IFAD, the Rockefeller Foundation and BMZ/GTZ, is recommending evaluation techniques that support learning and change and are fully consistent with those we have outlined in this paper (for example Herweg and Steiner, 2002ab; OECD,2002).

References

Adesina, A. A. and Chianu, J., 2002. Chapter 4: Farmers' use and adaptation of alley farming in Nigeria. In: Barrett, C. B., Place, F., and Aboud, A.A., (Editors), 2002. Natural resources management in African agriculture: Understanding and improving current practices. CABI Publishing in Association with the International Centre for Research in Agroforestry.51- 64.

Ashby, J. 2001. Integrated Research on Food and Environment: An Exit Strategy from the Rational Fool Syndrome in Agricultural Science. in: Campbell, B. M., and Sayer, J., A., (Editors) 2003. Integrated natural resource management: Linking productivity, the environment and development. CABI publishing UK in Association with Center for International Forestry Research (CIFOR) 293-310.

Ashby, J., 2003. Chapter 1. Introduction: Uniting Science and Participation in the process of innovation – Research for Development. In: Pound, B., Snapp, S., McDougall, C., and Braun, A., (Eds). Managing natural resources for sustainable livelihoods: Uniting science and participation. IDRC and Earthscan Books, London, UK. 1-15.

Barrett, C. B., Place, F., and Aboud, A.A., (Editors), 2002. Natural resources management in African agriculture: Understanding and improving current practices. CABI Publishing in Association with the International Centre for Research in Agroforestry. 335pp

Campbell, B. M., and Sayer, J., A., (Editors) 2003. Integrated natural resource management: Linking productivity, the environment and development. CABI publishing UK in Association with Center for International Forestry Research (CIFOR).315pp.

Campbell, B., Sayer, J.A., Frost, P., Vermeulen, S., Ruiz-Perez, M., Cunningham, A., and Ravi, P. 2001. Assessing the performance of Natural Resource Systems. Conservation Ecology 5(2): 22. [online] URL: *http://www.consecol.org/vol5/iss2/art22/index.html*

Ceccarelli, S., Grando, S., Tutilwer, R., Baha, J., Martini, A., Salahieh, Goodchild, A., Michael, M., 2002. A methodological study on participatory barley breeding. I. Selection phase. Euphytica 111: 91-104.

CGIAR. 2000. Integrated natural resource management research in the CGIAR. A brief report on the INRM (Integrated Natural Resource Management) workshop held in Penang, Malaysia, 21–25 August 2000. Consultative Group on International Agricultural Research Secretariat.

CIMMYT, SPIA, 2003. Why has impact assessment research not made more of a difference? Summary proceedings of an international conference on impact of agricultural research and development, held 4-7 February 2002 in San José, Costa Rica, and organized by the CGIAR's Standing Panel on Impact Assessment (SPIA) of the Interim Science Council and Economics Program of the International Maize and Wheat Improvement Centre (CIMMYT), CIMMYT, Mexico. DFID. 2003. Agriculture and poverty reduction: Unlocking the potential. A United Kingdom Department for International Development (DFID) Policy Paper, December 2003. 12pp.

Douthwaite, B., Delve, R., Ekboir, J., Twomlow, S.J., 2003. Contending with Complexity: The Role of Evaluation in Implementing Sustainable Natural Resource Management. International Journal of Agricultural Sustainability 1:51-68

Douthwaite, B., Ekboir, J., Twomlow, S., Keatinge, J. D. H, 2004. The concept of INRM and implications for developing evaluation methods. In: B. Shiferaw, H. A. Freeman and S. M. Swinton (Eds) Natural Resource Management in Agriculture: Methods for Assessing the Economic and Environmental Impacts of Management Practices. CABI (in press)

Ekboir, J. 2003. Why impact analysis should not be used for research evaluation and what the alternatives are. Agricultural Systems, 78. 166-184.

Ellis-Jones, J., Zvarevashe, V., Twomlow, S., Stevenson, K., 2001. Increasing productivity and improving livelihoods in semi-arid Zimbabwe. Research development linkages, a case study from southern Zimbabwe. Landwards, Spring 20001, 56 (1), 24-29.

Guijt, I., 1998. Participatory monitoring and impact assessment of sustainable agriculture initiatives: an introduction o the key elements. SARL discussion paper No. 1, July 1998. IIED, London. 112pp.

Hall, A., Rasheed Sulaiman, V., Clark, N., Yoganand, B., 2003. From measuring impact to learning institutional lessons: an innovation systems perspective on improving the management of international agricultural research. Agricultural Systems, 78: 213-241.

Harwood, R.R. and Kassam, A.H. (Eds). Research towards Integrated Natural Resources Management: Examples of research problems, approaches and partnerships in action in the CGIAR. Interim Science Council, Consultative Group on International Agricultural Research. Washington, DC, USA: 168pp.

Herweg, K., and Steiner, K. (Editors) 2002. Impact monitoring and assessment: Instruments for use in rural development projects with a focus on sustainable land management. Volume 1: Procedure. Volume 2: Toolbox. Centre for Development and Environment (CDE, Switzerland) and Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ, Germany). Volume 1 48 pp, Volume 2 44 pp.

Horton, D. and Mackay, R., 2003. Using evaluation to enhance institutional learning and change: recent experiences with agricultural research and development. Agricultural Systems, 78: 127-142.

Johnson, N., N. Lilja and J. A. Ashby. 2000. Characterizing and measuring the effects of incorporating stakeholder participation in natural resource management research: Analysis of research benefits and costs in three case studies. Working Document No. 17. PRGA Program. Cali, Colombia. 132 pp.

Kelley, T. and Gregersen, H., . 2003. NRM impact assessment in the CGIAR: meeting the challenge and implications for ICRISAT. In Shiferaw, B. and Freeman, H. A. (Eds). Methods for assessing the impacts of natural resource management research. A summary of the proceedings of an International Workshop, 6-7 December 2002, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India. Order code CPE148. 105-115.

Kerr, J., and Chung K., 2001. Evaluating watershed management projects. CAPRI Working Paper No. 17. Washington DC, USA. International Rood Policy Research Institute.

McCown, R.C. 2001. Farming systems research and practice. Proceedings of the 10th Australian Agronomy Conference, Hobart, Australia

McDougall, C, and Braun, A., 2003. Chapter 2: Navigating complexity, diversity and dynamism: Reflections on research and for natural resource management. In: Pound, B., Snapp, S., McDougall, C., and Braun, A., (Eds). Managing natural resources for sustainable livelihoods: Uniting science and participation. IDRC and Earthscan Books, London, UK. 20-47

Mitchell R., B Agle and Wood, D. 1997. Towards theory of stakeholder identification and salience: defining principle of who and what really counts. Academy of Management Review 22(4), pp. 853-886.

Muwira, K., Hagmann, J., and Chuma, E., 2001. Mainstreaming participatory approaches in SWC. Chapter 28 In: Reij, C., and Waters-Bayer, A., (Eds). Farmer Innovation in Africa. A source of inspiration for development. Earthscan publications Ltd., London. 310-324.

OECD, 2002. Sustainable Development Strategies. A resource book. Organization for Economic Co-operation and Development, Paris and United Nations Development Program, New York. Earthscan Publications Ltd. 358pp.

Perez, C. and Tschinkel, H., 2003. Improving watershed management in developing countries: A framework for prioritizing sites and practices. AgREN Network Paper No. 129, July 2003. ODI, London. 15pp.

Pound, B., Snapp, S., McDougall, C., and Braun, A., (Eds). Managing natural resources for sustainable livelihoods: Uniting science and participation. IDRC and Earthscan Books, London, UK

Pretty, J. and Hine, R., 2001. Reducing food poverty with sustainable agriculture: A summary of evidence. Final report from SAFE-World Research Project, University of Essex, Colchester, UK. Reddy, V. R., Soussan, J., 2003. Assessing the impact of participatory watershed development: a sustainable rural livelihoods approach. In: In Shiferaw, B. and Freeman, H. A. (Eds). Methods for assessing the impacts of natural resource management research. A summary of the proceedings of an International Workshop, 6-7 December 2002, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India. Order code CPE148.84- 93.

Rusike, J., S. S. Snapp, and S. J. Twomlow. 2004. " The Mother-Baby Trial approach for Developing Practical Soil Water Fertility and Management Technologies." Sourcebook on Participatory Research and Development. Manila: International Potato Center. Chapter 34,Sayer, J., A., and Campbell, B. M. 2003. Research to integrate productivity enhancement, environmental protection, and human development. Chapter 1 In: Campbell, B. M., and Sayer, J., A., (Editors) 2003. Integrated natural resource management: Linking productivity, the environment and development. CABI publishing UK in Association with Center for International Forestry Research (CIFOR). 1-14.

Shiferaw, B. and Freeman, H. A. (Eds). 2003. Methods for assessing the impacts of natural resource management research. A summary of the proceedings of an International Workshop, 6-7 December 2002, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India. Order code CPE148.

Snapp, S. and Heong, K. L. Chapter 4. Scaling up and out. In: Pound, B., Snapp, S., McDougall, C., and Braun, A., (Eds). Managing natural resources for sustainable livelihoods: Uniting science and participation. In: Pound, B., Snapp, S., McDougall, C., and Braun, A., (Eds). Managing natural resources for sustainable livelihoods: Uniting science and participation. IDRC and Earthscan Books, London, UK IDRC and Earthscan Books, London, UK. 67-83.

Twomlow, S. J. 2003. Soil and Water: The Flesh and blood of semi-arid agriculture in Africa. In: Waliyar, F., Collette, L., and Kenmore, P. E., 2002. Beyond the Gene Horizon: sustaining agricultural productivity and enhancing livelihoods through optimization of crop and crop-associated biodiversity with an emphasis on the semi-arid tropical agroecosystems. Summary proceedings of a workshop, 23-25 September 2002. Patancheru, India. Patancheru 502 324, Andhra Pradesh, Teme di Caracalla, Rome 00100, Italy: Food and Agriculture Organization of the United Nations (FAO). 36-65.