Research Into Use

sharing lessons to enable innovation in agriculture

Evaluation of Innovation Systems and Agricultural Research Programmes

Literature Review

DRAFT

2 August 2010

DFID Department for International Development

TANGO INTERNATIONAL
TECHNICAL ASSISTANCE to N.G.O.'S

TABLE OF CONTENTS

ACKNOWLEDGMENT

This document is an output from the Research Into Use Programme (RIU) funded by the UK's Department for International Development (DFID) for the benefit of developing countries. The views expressed are not necessarily those of DFID.

TABLE OF TABLES

INTRODUCTION

Research Into Use (RIU) is a programme of UK's Department for International Development (DFID) that began in 2006 and was intended to run until 2011. In 2008 a mid-term review of RIU recommended more focus in vision and more decentralisation in management. A following technical review further recommended RIU change its position to explicitly be a research project. As a result of both these recommendations, RIU developed a new business plan, DFID-approved, recruited a new team, including Ian Maudlin, and importantly, is now focussed on putting research into use for innovation by aiding learning and capacity building (RIU, 2010).

RIU will be evaluating six African Country Programmes in Rwanda, Tanzania, Malawi, Zambia, Nigeria and Sierra Leone as well as the Asia Innovation Challenge Fund of 13 projects in South Asia, four based in Bangladesh, five in India and four in Nepal. RIU's objective is the innovative research on the impact of different approaches that best use agricultural research to affect policy in development (ibid.).

To organize this research, RIU has established six "narratives": poor user-led innovation, public private partnership-led innovation, capacity development-led innovation, opportunity-led innovation, investment-led innovation, and research communication-led innovation. RIU has identified *Poverty Impact* (relevance, income, sustainability, etc.), *Social Impact* (targeting, gender, inclusion, exclusion, etc.), *Process Impact* (efficiency, effectiveness, etc.), and *Policy Impact* (influence, attribution, etc.) as possible central themes for its proposed impact evaluation.

Therefore, the purpose of this present literature review is to draw useful insights from research methodologies used in previous evaluations of Innovation Systems and Agricultural Research Programmes so that RIU may develop an evaluation approach grounded in these approaches. Relevant literature was solicited and gathered by TheIDLgroup Ltd and Technical Assistance to NGOs (TANGO) International and added to literature of which RIU was already aware. The review was guided by constraints of time and relevance.

This resulting review is organized into an overview of general information regarding impact evaluations, followed by potential evaluation approaches / methodologies, and concluded with key issues / lessons learned that will inform RIU's evaluation approach.

OVERVIEW

Given this background, and before describing those various approaches and methods, an overview of what is meant by "impact evaluation" with its purpose, trends, and criticisms will follow.

What is Impact Evaluation?

It's advisable that this overview begin with the definition of what it reviews in the literature, namely impact evaluation. There is interchangeable use of the terms "evaluation" and "assessment" in some literature while there is clearly distinguished denotation given in others. In this review, for the sake of gathering the broadest applicability to the evaluation of innovations systems and applied agricultural research programs, both "evaluations" and "assessments" are treated similarly and the term "evaluation" used preferentially.

To define "evaluation", the Active Learning Network for Accountability and Performance (ALNAP) and Channel Research Course Reference Manual (2010) cites the Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC) glossary definition that—abbreviated—defines "evaluation" as *(1) the assessment of a programme to determine its fulfillment of objectives and provide useful information for the decision makers, (2) the relatively objective process of determining the worth of a development intervention* (cf. ALNAP 2010 and OECD/DAC 2002).

To define "impact", White (2009a) clarifies in 'Some Reflections on Current Debates in impact evaluation' that there are two equally valid but mutually exclusive definitions: impact defined generally as *any outcome of an intervention* verses impact defined as *the difference between the outcome in the presence of the intervention against the outcome in absence of the intervention*. Jones, Jones, Steer & Datta's (2009) impact evaluation follows the second definition as an assessment of whether or not a development intervention is having the appropriate effect by comparing the impact of that intervention against outcomes in the absence of that intervention.

Impact evaluation designs have been both experimental—by establishing the intervention in a random definable group and comparing outcomes with those of a definable group where the intervention is not established—and quasi-experimental—in the absence of randomness and replacing the control with statistical simulation (ibid.). Karlan sees randomization as improving reliability when possible (Chambers, Karlan, Ravallion & Rogers, 2009).

What Evaluation Does

There are three reasons for impact evaluations outlined by Karlan: to determine where limited resources should be invested, to determine how programmes can be improved, and to encourage continued investment (Chambers, Karlan, Ravallion & Rogers, 2009).

In addition to this rationale, impact evaluations also add certain potential benefits, listed by the European Initiative for Agricultural Research for Development (EIARD, 2006): (1) increased impact of investments in development; (2) creation of learning opportunities; (3) informing the management, priorities, planning of the research system; (4) enhanced advocacy and funding opportunities with information about the impacts.

Impact evaluations may also provide learning opportunities for others in the international development field. By engaging in thorough evaluation of innovation within agricultural extension systems, researchers can promote a better understanding of institutional relationships, the measurement of research outcomes, challenges to knowledge dissemination, and the ways in which they may be overcome. Ultimately, this knowledge can help researchers improve their own performance and enhance their contribution to ongoing efforts toward reducing poverty, alleviating food insecurity, protecting the natural environment, demonstrating the results of the intervention (Horton & Mackay, 2003), as well as increasing incomes.

Finally, in an assessment of organisational learning with respect to agricultural research, International Food Policy Research Institute (IFPRI) found that Social Network Analysis (SNA)—which considers the size, efficiency, and connectedness of various actors within a particular social network—was useful for mapping technical, commercial and information flows (Ekboir et al., 2009).

Evaluating Innovation Systems

Beyond the reasons for using impact evaluations, there are challenges involved in evaluating innovation. Review of evaluations assessing innovation in agriculture extension (e.g. Farmer Field School) revealed lack of agreed conceptual frameworks for carrying out such studies (Waddington, Snilsveit, White & Anderson 2010). EIARD (2006) argues that research does not sufficiently address the needs of the stakeholders and that assessments and evaluations should be designed to increase the chances that research will benefit the impoverished, communicate return of investment to donors, inform future investments, and educate the public. Patel (2007) calls for the models of research evaluations to be reconsidered. Revised principles should:

- create uninterrupted dialogue among agencies through which to share perspectives on strategies, results, and opportunity;
- address the risk-taking nature of funding with the implications on innovation and precedent;
- discuss clearly the opportunity assessment presents for measuring the impact of scale, its production, sustainability, and drivers;
- build broader support for more diverse strategies and roles; and
- create systems for incorporating anecdotal information from people in the targeted communities.

Institutional Lack of Social Knowledge. According to Guijt (2007), social change has interlinking features that have significant affect on assessments, namely, it is non-linear as well as unpredictable, takes a lot of effort from a lot of fronts, has "fuzzy boundaries", has results that are difficult to recognize and is long term. International agricultural research agencies (e.g. CGIAR* members) work very closely with national partners; as such their international programmes play a strategic role in disseminating research findings, methodologies, and institutional strategies that national programmes can use to generate finished technologies for farmers. Hence, the impact of international programmes should be assessed primarily in institutional terms, not in terms of production increases at the farm level (Horton, 1986).

However, past evaluations have revealed that academic researchers must have a basic understanding of the social networks within the operating environment in order to translate knowledge into improved practices (Ekboir, et al. 2009; cf. White 2009b). Accordingly, evaluation methodologies should assess the extent to which participating research institutions possess this knowledge and incorporate it into innovative technologies. In this regard Gujit states that the following critical considerations must be taken into account when designing appropriate evaluation systems for measuring social change (ibid.):

- Evaluation design should reflect an understanding of social change (including evolving development policies and strategies) and underlying assumptions regarding its causes;
- Caution should be taken in determining 'causality' or attributing credit for social change when evaluating a 'multi-actor, multi-location, multi-level, multi-strategy' programme;
- Evaluations must make the most of available indicators and understand their limits;

* CGIAR is an acronym for the Consultative Group on International Agricultural Research. CGIAR has adopted a consortium arrangement to support the research and dissemination of findings from 15 international research centers. Detail on each of the centers can be found at: http://www.cgiar.org/centers/index.html

- Evaluations of social change should ensure the capacity to facilitate critical reflections on power, justice, and policy processes as well as the ability to engage in learning processes; and
- Evaluations should establish clear and ethical standards that appropriately address unequal relations between North and South, donors and grantees, external experts and local people.

Quantitative vs. Qualitative. Jones, Jones, Steer & Datta (2009) note that quantitative methods are given widespread hierarchical preference in the recognition of scientific rigor over qualitative methods, citing the World Bank (IEG, 2006) and the Center for Global Development (CGD 2006). White (2009a) argues that no hierarchy exists, only that many contexts are best served by quantitative methods while there are also many other contexts best served by qualitative methods. Despite his own view that quantitative methods are more often than not the best, he cautions that the best methods for the context must be adopted.

Perrin (2000) considers qualitative methods of evaluation to be appropriate when used in combination with quantitative methods. Where there are trade-offs involved in a new technology, such as between average productivity increases and higher vulnerability, qualitative research methods can improve understanding of how different categories of households and individuals value those trade-offs. Using focus groups and other qualitative methods is also useful in identifying factors that might otherwise be overlooked, or to prioritize which of the many potential effects are important for poor people in that particular area (Adato & Meinzen-Dick, 2002).

In consideration that there is much less literature employing the more rigorous definition of "impact evaluation", White (2009a) calls for more experimental designs (as opposed to quasi-experimental). Research suggests there has been relatively little detailed analysis regarding the impact of technological innovation on local production systems, particularly in marginal and semi-commercial areas. Many assessments rely on experimental results and economic models rather than empirical studies on how farmers' production systems have changed with the introduction of the new technology (Horton 1986). In many cases , evaluations of agricultural extension activities are either: (1) designed to be statistically rigorous, but with limited scope, or (2) designed to be more comprehensive and detailed in terms of subject matter, but with limited coverage (Waddington, Snilsveit, White & Anderson, 2010).

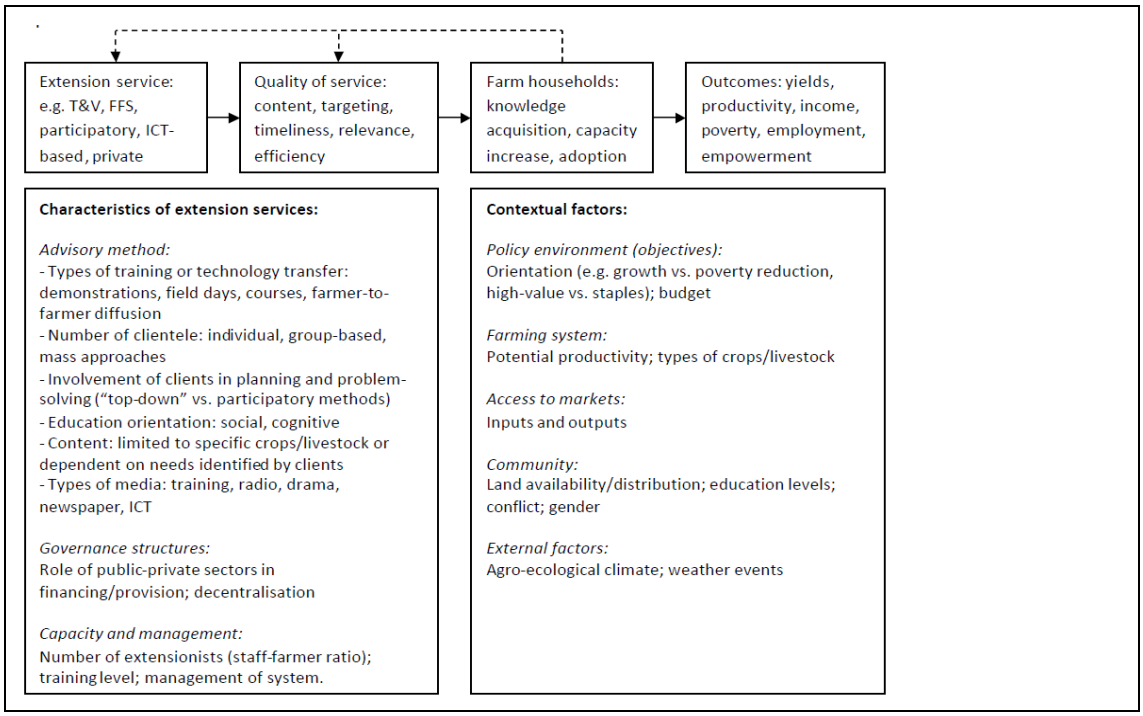
Controls & Counterfactuals. Related to the challenges of scientific rigour is the question of a counterfactual, control group. Jones, Jones, Steer & Datta's (2009) require a counterfactual in the impact evaluation. However, White (2009a) argues that a control/counterfactual is in many cases unneeded because it is implicit. He surveys the options for measuring the outcome *sans* intervention. These include measuring the community before the intervention, which is common in impact evaluations, but unreliable, White maintains, because of limitless potential variables (ibid.).

Indicators. Another challenge for evaluations is choosing those elements by which to measure impact. Evaluation approaches for innovation systems need to be flexible enough to identify and respond to the exploration and creative energy that is an inherent asset for such systems. This requires a set of indicators that is inclusive enough to capture unexpected results and valuable lessons regarding both processes and outcomes (Ekboir et al., 2009).

In Table 1 Birner et al. (2006) provide a causal model below for understanding the influence of certain characteristics on the uptake and effectiveness of agricultural extension services. The causal model is intended to identify the basic elements that influence the effectiveness of such services—from extension service inputs, to indicators of service quality, farmer adoption of improved

technologies/practices, and contextual factors influencing the acquisition of knowledge. Review of the framework identifies several potential indicators that could be used in the present study.

Table 1: Birner et al. (2006) Causal model and characteristics of services and… conditions….



Source: Birner et al., 2006

This model (ibid.) implies that the effectiveness of extension systems in building capacity, promoting adoption of improved technologies and supporting improve agricultural outcomes, depends on key factors including the advisory methods used, prevailing governance and extension management structures, and underlying contextual factors (policy environment, market access, characteristics of beneficiary communities and weather conditions).

In a critical review of literature on the impact of different approaches to extension, Anderson (2007) considered the influence of various governance structures, approaches to capacity and management, and advisory methods. The review highlights a general lack of verifiable information on the impact and cost-effectiveness of agricultural innovation. Anderson concludes that "the existing studies do not make it possible to identify which of those reform elements is effective under which circumstances" (26 Anderson, 2007).

A Birkhaeuser et al. (1991) comparison found that most evaluation studies rely on survey data from one point in time and typically use multivariate (limited dependent variable) estimation techniques, with inadequate control for sample selection bias. Some researchers warn against the use of time-series production data for measuring research impact given the traditional inaccuracy of production estimates at the national level and the numerous factors that influence changes in production at the sub-national level (Horton 1986) as well as attribution issues.

Perrin (2000) alleged that most evaluations fail at supporting innovation because their indicators are changes in the mean rather than the few demonstrable cases of meaningful innovation. Innovation being infrequent and unpredictable, an approach to impact evaluations is needed that accounts for a

few gainful innovations in the midst of many more failed innovations (ibid.). Perrin champions the use of venture capitalists' point of view in impact evaluations. Evaluations should consider the overwhelming majority of interventions as learning opportunities with which to identify best practices in the process of eventual innovation. Evaluations should not penalize interventions for taking a risk.

Several other sources have demonstrated indicators that proved effective. The integration of economic analysis and social network analysis (SNA) through the IFPRI study identified the following indicators of efficiency and effectiveness of information exchange at the local level (Conley & Udry, 2001; Krishnan & Sciubba, 2004; Matuschke 2008; Munshi 2004):

- Homogeneity of groups (information sharing is more efficient among homogenous groups)
- Similarity between practitioners and researchers (farmers respond more readily to individuals who are similar to them)
- Number of nodes and social ties between practitioners (information sharing is more efficient in areas characterized by a social cohesion)

Horton (1986) describes other work that aimed at evaluating international agricultural research and development programmes. He recommends that two distinct types of impact be assessed: production impact and institutional impact. "Production impact" refers to the physical, social, and economic effects of new cultivation and post-harvest methods on crop and livestock production, distribution, and use and on welfare in general (including the effects on employment, nutrition, and income distribution). "Institutional impact" refers to the effects of research and development (R&D) systems on the capacity of research and extension programmes to generate and disseminate new production technology (ibid.).

Requirements for a comprehensive assessment of the production impact of agricultural research include: (1) estimates of changes in production in regions or countries over time; (2) knowledge of major production systems, including "old" and "new" technologies, and the proportion of total output generated by each; and (3) knowledge of production functions for each system. If distributional effects of technological change are to be analyzed, estimates of supply and demand elasticity are also required (ibid.).

EIARD (2006) has also recommended that assessments address the "intermediate processes" such as the perceptions of stakeholders, the opinions of the community as to quality and relevance, and changes in the behaviour of individuals and groups.

Bias. How research bias can influence impact evaluation is another challenge. There are three common types of bias in agricultural research programmes. These research biases are well known but the analyses used in most evaluations do not allow for their control (Romani, 2003; Waddington, Snilsveit, White & Anderson 2010):

- Endogenous placement bias may occur where programmes are situated in areas seen as more likely to be receptive to extension services;
- Selection bias occurs where skilled and knowledgeable farmers are more likely to seek out extension services, and although this source of bias may be reduced if extension agents initiate contact with the farmers, agents themselves may also rather work with more experienced farmers; and

- Simultaneity bias arises in the sample of farmers visited by extension services if farmers only contact extension agents when they have problems.

For instance, on-farm research is critical for evaluation of technological innovation, but is prone to bias. Researchers tend to favour small plots (more time, attention, inputs, capital) in an effort to show results, locate research sites on better than average land, fail to incorporate control groups into evaluation studies, and employ logical approaches not followed by target farmers (Horton 1986).

Participation. One final potential challenge is that of inclusion of stakeholders in the impact evaluation. Guiit (2007) recommends that evaluation of programmes like RIU must adopt a more participatory approach to learning that increases accountability to the end-users of information at the grass-roots level. Participation in assessment by actors/stakeholders can benefit them by (1) structuring the way they gather knowledge, (2) increasing their network of interaction, (3) enhancing communication between them, (4) increasing their investment in the project, and (5) boosting their morale (Horton and Mackay, 2003). Chambers (Chambers, Karlan, Ravallion & Rogers, 2009) sets the priority for Participatory Methods to be securing "good facilitator innovators" (p 4).

Evaluating Policy and Research

In addition to the challenges involved in evaluating innovation, there are also challenges to evaluations influencing policy. Echoing Guijt's (2007) observations of social change, Young and Mendizabal (2009) observe that "policy processes are complex and rarely linear or logical" (p 1). This note leads the six lessons they share from their five years working at influencing policy, which also include:

- "many policy processes are only weakly informed by research-based evidence"
- "research-based evidence can contribute to policies that have a dramatic impact on lives"
- "policy entrepreneurs need a holistic understanding of the context in which they are working"
- "policy entrepreneurs need additional skills to influence policy"
- "policy entrepreneurs need clear intent—they need to really want to do it" (pp 1-2).

The work done by the Overseas Development Institute's Research and Policy in Development (RAPID) and the IDRC on Outcome Mapping is key here. (The latter will be discussed in "Outcome Mapping / RAPID Outcome Mapping"). Their work is also about measuring social change. The RAPID programme's aim is to improve how research as well as evidence is used in influencing policy and change (Start.& Hovland, 2004). They focus their work in four areas: using evidence to identify and develop policy, improving systems in the development field, managing knowledge better to enhance impact, and promoting evidence-based policy (ibid.).

In the paper 'Making a difference: M&E of policy research' (2007), Hovland surveys the full range of approaches currently taken in policy research. These include those approaches that evaluate strategy and direction (logframes, impact pathways, modular matrices, and social network analysis), approaches that evaluate management ('fit for purpose' reviews, appreciative inquiry, 'lighter touch' quality audits, and horizontal evaluation,), approaches that evaluate outputs (after action reviews, evaluating networks, evaluating websites, evaluating policy and briefing papers, evaluating academic articles and research reports), approaches that evaluate uptake (impact logs, user surveys, new areas for citation

analysis), and approaches that evaluate outcomes and impacts (outcome mapping, episode studies, most significant change, innovation histories, rapid outcome assessment; p 3).

POTENTIAL EVALUATION APPROACHES / METHODOLOGIES

In this section different approaches or methods to impact evaluations are presented. These have been extracted from the literature reviewed and given the same relative attention given in those sources. The literature itself seems to move from mainstream approaches based on accountability toward mixed, alternative, and innovative approaches, this section is organized similarly. This section begins with what the literature has to say about the overall approach to methods in evaluations and assessments.

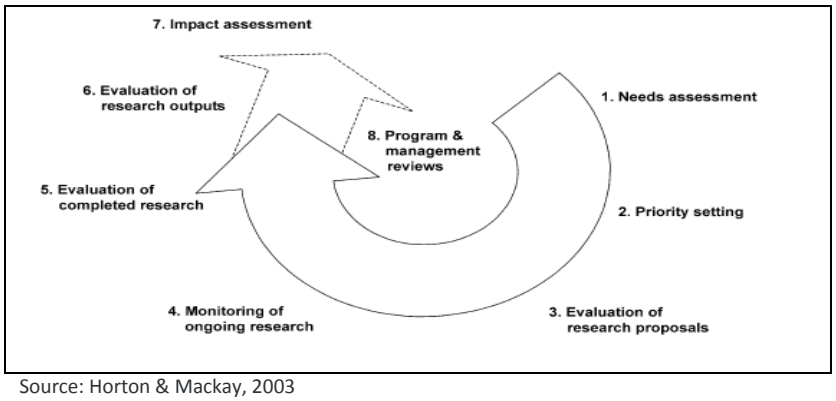
Overall Evaluation Approaches

Demonstrating the breadth and diversity of approaches to evaluations and assessments overall, the categories in Guijt (2007), Horton and Mackay (2003), and Kristjanson et al. (2002) are especially helpful.

Impact evaluation approaches are informed by different perspectives on social change assessments. Guijt (2007) categorizes these perspectives, “action research/appreciative inquiry” includes those seeking to simultaneously understand the social structures and affect social change. “Organizational learning” perspectives includes those that promote learning as a part of the intervention. “Popular education” approaches relate individuals’ experience to broader social issues in an attempt to increase their influence. “Feminist evaluation” uses the knowledge resulting from assessments to work toward community emancipation. Although similar in nature, the “participatory/empowerment evaluation” approach emphasizes the community assessing, improving, and strengthening their own efficacy. “Democratic evaluation/dialogue” approaches address social issues by increasing accountability and transparency. Finally, there is the approach of “Utilization-focused evaluation” that evaluates for the sake of learning itself, relevant to social change.

In addition to being informed by different perspectives, evaluation methods differ at the different stages in the research process that they can be and often are utilized by research institutions. As Table 2 illustrates, Horton and Mackay (2003) categorize evaluation methods by the stage in which they occur.

Table 2: Horton (1998) Eight types of agricultural research evaluation



Source: Horton & Mackay, 2003

The first four types of evaluation identified in the framework are generally carried out to support internal decision making. The second four—evaluation of completed projects, evaluation of research

outputs, impact assessment and programme reviews—are typically conducted to meet external accountability requirements. Their primary purpose has not been to help researchers better understand the results their activities produce in the field, nor to improve the process of information dissemination (Horton & Mackay, 2003).

Categorisable by stage and perspective, evaluation methods can also be categorised by the means with which they approach the target object and context. Table 3, from the Kristjanson et al. (2002) review of evaluations, summarizes methods by approach/examples, uses/strengths, and weaknesses.

Table 3: Kristjanson et al. (2002) Summary of evaluation approaches

Approach and examples	Uses and Strengths	Weaknesses
Village workshops/ discussions, stakeholder consultations, key informant interviews (Kristjanson et al., 2002b, Franzel et al., 2002)	Good for identifying key impacts and indicators and to identify key factors affecting adoption, diffusion and impact; institutionalizing impact assessment in a village	Community leaders can dominate discussions, women are often left out, participants tell the organizers 'what they want to hear'. Lessons may not be applicable across broader areas due to unique characteristics of villages chosen
Community-level formal surveys (Pender and Scherr, 1999; Okike et al., 2001)	Useful for gaining an understanding of the characteristics of communities that are benefiting from a new technology and village-level factors affecting adoption and diffusion of new interventions	Quality of information can be dependent on relatively few individuals and may vary considerably across communities
Household-level formal surveys for looking at adoption and impact (Adesina et al., 2000; Baidu-Forson, 1999; Nicholson et al., 1999)	Possibilities for studying: household-level characteristics influencing adoption and impact; farmer knowledge / understanding of new strategies; asset changes; food/nutrition/health impacts; income/expenditure impacts; social/cultural changes; labour/farming strategy changes	Time consuming and relatively expensive; results often not available for some time (and often not communicated back to the participants). Typically targeted towards household heads, whereas resources may be controlled by different household members. Many sensitive and often difficult questions subject to cross-cultural bias. Tend to treat adoption as an event rather than a process of learning/ experimenting; most beneficial with a well-defined technology that has been in use for a long time (ex post studies)
Financial and economic analyses of the production effects of new technologies (Place et al., 2002)	Monitoring of labour or other resource requirements for a particular integrated crop-livestock system. Analyses of costs/benefits from researcher- and farmer-designed experiments	There will be benefits and/or costs that are not captured in the marketplace (e.g. bank account or insurance aspect of cattle)
Transect walks, aerial photography (Reid et al., 1997; N. de Haan, pers. comm.)	Can help estimate the numbers and locations of users/adopters of crop-livestock interventions, can measure stocks of broad-scale natural resources	Limited area coverage, getting cheaper but still can be quite costly; rarely useful for measuring flows of natural resources
Spatial analysis; GIS; satellite imagery (Staal et al., 2002)	Allows forward-looking approach to understanding systems and impacts of things like population growth, climate change. Also allows delineation of target zones or recommendation domains for specific technologies	Dependent on good spatial datasets; easier to identify crop-based systems than livestock; little ability to accurately distinguish smallholder systems over broad scales
Plot and landscape field measurements of natural resources; both stock and flows	Allows accurate measurement of impacts of policy and management interventions; forms basis for improvement of process-based models; community participation should be key element	Often are avoided because of cost, but no substitute is yet available; process-level information is long term and quite expensive

Table 3: Kristjanson et al. (2002) Summary of evaluation approaches

Approach and examples	Uses and Strengths	Weaknesses
Human perceptions of environmental change (household/community surveys	Often one of the only ways to estimate environmental change over time; good for hypothesis development	Restricted by human abilities to sense environmental change
Market studies (Scarpa et al., 2002; Turner and Williams, 2002; Fafchamps and Gavian, 1997)	For analysis of differences in market conditions; demand studies, determinants of prices and spatial integration of markets	If one-shot surveys, conditional on weather or other circumstances particular to timing of survey; solid time series analysis requires regular market visits over a long period
Economic surplus methods (Alston et al., 1995)	For investigating the effects of interventions that have measurable impact on the production and price of commodities; both ex ante and ex post	Requires good information on price responsiveness of producers and consumers that often just is not available; non-marketed benefits and hidden costs (e.g. social) difficult to incorporate
In-depth anthropological/ sociological and characterization studies; farmer assessments (Ashby, 1990)	Good for characterizing major household types; identification of important livelihood strategies of households and pathways for poverty alleviation; processes of testing and uptake of new technologies; more detailed knowledge on resource allocation and investment	Time consuming, relatively expensive and cover a small number of households/area

Source: Kristjanson et al., 2002

Though these papers demonstrate the scope and applicability of different evaluation methods, there remains the concern for efficacy.

Results-Based Management

In response to an increasing demand for accountability, public and private sector organisations have increasingly adopted results-based management which attempts to link inputs, outputs, outcomes and impacts. However, previous studies have shown that while results-based management is an improvement over conventional compliance evaluation in assessing the effectiveness of programme activities, it is not an appropriate methodology for assessing outcomes (Kusek & Rist, 2004; Pestieau, 2003; Peterson, Gijsbers & Wilks 2003).

Speaking to this, Guijt (2007) states the documented inadequacy of mainstream approaches like many of those listed above and refers to the findings of the Gray Rocks conference of the Assessing Social Change (ASC) group. They determined that despite such inadequacy, many organisations continue to use these mainstream approaches. This is often in response to the need to show measurable and attributable change. However, mainstream monitoring approaches have proven ineffective in capturing the impact of efforts aimed at capacity building, social change, promoting innovation and enhancing information networks (ibid.).

In regard to the adequacy or inadequacy of various methods, descriptions of Social Innovation Assessment, Utilization-Focused Evaluations, Holistic Livelihoods Methods, Sustainable Livelihoods Methods, Rapid Assessment, and Mixed-Method/Alternative/Innovative Analysis was found to be instructive.

Social Innovation Assessment

One approach, termed Social Innovation Assessment (SIA), looks to combine both performance and impact assessment in evaluating the effectiveness of social innovation systems supporting agricultural research. SIA follows a phased approach to evaluation whereby development of a social innovation survey questionnaire is preceded by intensive, qualitative case studies (Pant & Odame, 2008).

SIA highlights the inherent difficulty in attributing impact amid dynamic contextual factors and stakeholders. It argues in favour of evaluating the influence of a range of actors and factors that highlight important contributors and constraints to social innovation (ibid.).

Utilization-Focused Evaluation

Another approach entitled '*Utilization-Focused Evaluation*' (UFE) is focused at the micro-level (individual farmers, institutions) and the relationships that must be built between the evaluator of innovation systems and the end-user in order to determine research needs. Under UFE, end-users should play a direct role in determining the research focus and methods used for evaluating these systems (Patton & Horton 2009).

Sustainable Livelihoods

For plant breeders, soil scientists, and other technologists, the livelihood framework serves the purpose of linking their specific work and capacities with what people are capable of doing, what they are looking for, and how they perceive their needs. The livelihood framework thus provides a guide for research and intervention (Adato & Meinzen-Dick, 2002). However, using a sustainable livelihoods framework is not the simplest or most direct way of examining the impact of agricultural research. This is because it requires consideration of a wide range of counteracting factors at multiple levels, each of which may or may not have a significant impact on adoption of new technologies. But this reflects reality. Agricultural production activities may only be one component of a household's overall livelihood strategy. Understanding the other factors that influence the choice of livelihood strategies and specific technologies can be critical to improving the ultimate impact of agricultural research (ibid.).

Applying a sustainable livelihoods approach highlights the multilayered interactions between technologies and the vulnerability context of households – their asset base, access to social capital, and livelihood strategies. However, additional aspects of culture, power, and history need to be integrated with the framework to understand the role of agricultural research in the lives of the poor. Explicit attention must also be paid to the influence of gender, ethnicity, class and other types of social differentiation (ibid.).

EIARD (2006) argues that evaluations of agricultural research for development should be designed and carried out within a holistic livelihoods framework. While they acknowledge that measuring performance in terms of products developed and rates of return are important, they maintain that impact assessments must ultimately determine the extent to which research products and services are being used and how their use affects people's lives, their societies and environments.

Agriculture and Rural Development (ARD) (2006) identifies four analytical elements that should be carefully considered in evaluation of innovation systems: 1) key actors, 2) attitudes and practices, 3) patterns of interaction; and 4) enabling environment (see Table 4). These diagnostic elements should be assessed across developmental phases: the pre-planned (before interventions/opportunities), and foundation (when opportunities/support has been identified) in orchestrated systems, expansion (with

government intervention); nascent (when entrepreneurial identification of opportunities occurs), emergence (when supported by rapid growth by the private sector), and stagnation (from external pressures to integrate) in opportunity-driven systems.

Table 4: ARD (2006) Key analytical elements in evaluating dynamic innovation systems

Analytical element	Orchestrated systems			Opportunity-driven systems			
	Pre-planned	Foundation	Expansion	Nascent	Emergence	Stagnation	Innovation
Actors	Traditional public research organizations and private sector actors	Strong public sector presence; increasing private sector activity	Public, private, and civil society actors; emerging coordinating bodies	Private sector and/or civil society actors active	Primarily private sector actors	Most actors in place, but coordinating bodies still ineffective	Coordinating bodies well positioned to support all main actors
Attitudes and practices	Ivory tower mentality; limited trust	Traditional roles predominant	Willingness for collaboration	Opportunistic behavior (private sector)	Self-relying private sector	Un-coordinated, independent attempts at supporting the sector	Openness to partnering, collaboration, and inclusion
Patterns of interaction	Very limited interaction between main actors; limited access to information	Limited interaction between the main actors taking place	Interaction well developed within the clusters	Very limited networking	Informal private sector networks; poor contact with research	Collaboration weak	A dense network of interactions
Enabling environment	Generic research and training services available at most	Research and training services in place; limited incentives for private sector activity	Incentives for research, training, and private sector activity in place	Generic research and training services available	Incentives not in place; research, training, and financing sectors disconnected from the sector	Increasing incentives; research, training, and financing sectors still disconnected from the sector	Incentives and resources for research, training, and financial sector participation

Source: ARD, 2006

According to the authors, opportunity-driven systems typically take one of two 'trajectories' – orchestrated systems or opportunity-driven systems. In general, orchestrated innovation systems resemble those put in place by government-supported agricultural extension services. Alternatively, opportunity-driven innovation systems are typically supported by private interests in response to emerging or newly identified market opportunities (ibid.).

This framework was tested and found valid in case studies selected from sectors demonstrating: (1) strong growth in specialized markets, (2) strong integration in global market, (3) transformation by and therefore implications for food chain, (4) significant opportunity for employing the impoverished (ibid.).

Innovation Systems

Using the Innovation Systems framework, Hall, Mytelka, and Oyeyinka (2006) outline key elements for exploring agricultural innovation capacity in a rapid assessment by non-experts with limited training. The first aspect that Hall, Mytelka, and Oyeyinka address in their approach is the "sector timeline and evolution," which includes understanding the combination of triggers for new sectors and activity clusters. These include changes in policy or market, developmental interventions, and the sector's evolution of growth. Statistics that could inform analysis of sector timeline and evolution would include "value, size, growth rate, employment potential, nature of domestic and international market". These figures could be obtained by investigating secondary documentation, sector investment reviews, as well

as interviews with key informants and sector specialists. Hall, Mytelka, and Oyeyinka recommend triangulation of sources to balance the potential for competing or alternative perspectives.

The second element addressed by the authors is “sector mapping”. This step lists the actors and clarifies their roles, including their efficacy. It also includes identifying relevant organisations, the extent of their competency, their linkages and nature the of them (Hall, Mytelka, and Oyeyinka, 2006).

Arnold and Bell (2001) argue that relevant organisations can be classified according to their generation and/or use of knowledge: Under this system, organisations fall into the research domain (producing codified knowledge), the enterprise domain (consuming codified knowledge to produce tacit knowledge), the demand domain (consuming/marketing knowledge), and the intermediary domain (disseminating knowledge). Determining the classification of specific organisations can be done by investigating investment reviews, innovation policies and interviews with key informants and sector specialists. The specific competencies of individual organisations should also be delineated, such as management, research, marketing, etc. This information can also be gained from secondary sources, such as annual reports, and interviews with key informants and sector specialists. Note should also be taken of how each actor actively creates and develops additional opportunities as this indicates their responsiveness to external factors (Hall, Mytelka, and Oyeyinka, 2006).

Analysis of organisational linkages should begin by mapping general patterns of interaction followed by the nature of such relationships then the purpose of each. Hall, Mytelka, and Oyeyinka (2006) first recommend using a matrix that lists actors on both axes with notes on the specific interaction between each actor made at each connecting node. This tool is simple to use and makes obvious where interactions are absent. They then recommend a typology to classify the kinds of links existing between actors (see Table 5) and identify absent links that should be developed for innovation. Caution should be taken to be sure the “right types of linkages exist in the right place” (p 25).

Table 5: Hall, Mytelka & Oyeyinka (2006) Typology of partnerships and learning

Type of linkages	Purpose	Type of learning
Partnership	Joint problem solving, learning and innovation, may involve a formal contract of memorandum of understanding. Maybe less formal, such as participatory research. Highly interactive. May involve two organisations or more.	Mainly learning by interacting. Also learning by imitating and learning by searching.
Paternalistic	Delivery of goods, services and knowledge to consumers with little regard to their preferences and agendas	Learning by training
Contract purchase of technological or knowledge services	Learning or problem solving by buying knowledge from elsewhere. Governed by a formal contract. Interactive according to client contractor relations. Usually bilateral arrangement. Highly focused objectives defined by contract concerning access to goods and services.	Learning by imitating and mastering Might involve learning by training
Networks	Maybe informal or formal, but the main objective is to facilitate information flows. Provides know how and early warning information of market, technology and policy changes. Also builds social capital, confidence and trust and creates preparedness for change, lowering barriers to farming new linkages. Board objective	Learning by interacting Learning by searching
Advocacy linkages to policy process	Specific links through networks and sector associations to inform and influence policy.	Learning by interacting Learning by searching

Alliance	Collaboration in the marketing of products, sharing customer base, sharing of marketing infrastructure. Usually governed by a memorandum of understanding. Can involve one or more organisations. Board collaborative objective.	Learning by doing
Linkages to supply and input and output markets	Mainly informal but also formal arrangements connecting organisations to raw materials, inputs and output markets. Includes access to credit and grants from national and international bodies.	Limited opportunity for learning Some learning by interaction

Source: Hall, Mytelka & Oyeyinka, 2006

The third aspect Hall, Mytelka, and Oyeyinka (2006) address is the degree to which organisational habits and practices promote or restrict innovation. Broad habits should be outlined first, working toward those that are more subtle. This information could be gained from specific habit and practice studies of organisations, personal interviews, and interviews with key informants/sector specialists.

In the context of describing the habits and practices of organisations, they describe three subtle factors that distinguish institutions from one another in their relation to innovation. They include factors that affect: 1) interactions, sharing of knowledge, and learning; 2) risk taking and investments; and3) inclusiveness.

The fourth and final aspect they address is structures of policy and support. Specifically, they seek to determine: 1) whether or not such structures are effective in stimulating innovation; 2) whether or not they are relevant to or influenced by habits and practices; and 3)whether or not they should be adapted and if so, how. Important areas of inquiry should include the impact of policies and support on farmers as well as other actors, linkages, and local innovation systems. This information could be gained from relevant government representatives (ministry staff) and personal interviews with key informants and sector specialists (Hall, Mytelka & Oyeyinka, 2006).

Theory of Change / Theory-Based Evaluation

Theory of Change (TOC) is a well-established approach to evaluation. Connell and Kubisch define the TOC approach as "a systematic and cumulative study of the links between activities, outcomes and contexts of the initiative" (16 Connell & Kubisch, 1998). White (2009a,b) defines a theory-based approach as one that seeks to demonstrate the causal thread and thereby explain the 'why' of an invention's impact. It therefore has greater potential to affect policy. White finds that few examples realize this approach's potential.

Connell and Kubisch (1998) identify three stages to carrying out a TOC approach:

1. "surfacing and articulating a Theory of Change" (ibid.), meaning the generating of a theory by beginning at the desired final outcomes, laying out backward activities that affect those outcomes, diagramming required against existing resources then reconciling various perspectives (cf. White 2009b).

2. "measuring a CCI's activities and intended outcomes" (Connell & Kubisch, 1998), which should generally be taken to mean collecting data on the impact on the target object.[†] Connell and

[†] CCI is the acronym for Comprehensive Community Initiative or, sometimes, Community Change Initiatives. These are basically interventions focused on children, youth, and families that consider the broadest context, include all community sectors, extend long term, build collaborative relationships, with participatory management (cf. OJJDP, 2010).

Kubisch explain that measurement should be repeated and cumulative over early, intermediate, and later stages of the initiative. Establishing expected threshold levels of change is important in a TOC approach because there may not be an ex post analysis of whether change was statistically significant.

3. "analyzing and interpreting the results of an evaluation, including their implications for adjusting the initiative's theory of change and its allocation of resources" (Connell & Kubisch, 1998).

Connell and Kubisch (ibid.) identify a good TOC as being plausible (having activities that can plausibly lead to desired outcomes), doable (having available economic, technical, political, institutional, and human resources to carry out the intervention), and testable (being specific and complete enough to track credibly).

Drawing upon a variety of sources of information, including programme experience, scientifically generated knowledge, and community insights can contribute to success in using the TOC approach. A TOC approach can improve planning and implementation and facilitate the data collection during the evaluation process. This approach reduces problems associated with causal attribution of impact by making clear how activities are intended to lead to interim and longer-term outcomes, and examining contextual conditions that may affect activities (ibid.).

Perhaps the most important contribution of the TOC approach to the evaluation process is its emphasis on understanding. But to fully realize its potential, White says a theory-based evaluation must include (2009b):

1. An Embedded, Mapped Causal Chain, Reflecting Changes, Informed by Preceding Field Work,
2. An Understanding of the Social, Political, Economic Context Informed by Project Documents,
3. A Sample Size And Identification Of Sub-Groups Anticipate Differential Impacts,
4. Rigorous Factual Bivariate Analyses at Various Levels with a Credible Counterfactual,
5. Rigorous Factual Bivariate Analyses at Various Levels with a Credible Counterfactual,
6. Wide Range of Qualitative and Quantitative Methodologies Including Action Research.

Outcome Mapping / RAPID Outcome Mapping

Each publication on Outcome Mapping has its own unique spin on the steps involved. As initially developed by Canada's International Development Research Centre (IDRC), outcome mapping is based on the following principles:

1. "Actor-centred development and behavior change"
2. "Continuous learning and flexibility"
3. "Participation and accountability"
4. "Non-linearity and contribution, not attribution and control" (1-2 Jones & Hearn, 2009).

The steps associated with RAPID Outcome Mapping (ROMA) are:

1. “define a clear, overarching policy objective”
2. “map the policy context around that issue,” identifying key factors influencing policy,
3. “identify the key influential stakeholders”
4. “develop a theory of change”
5. “develop a strategy to achieve the milestone changes in the process”
6. “ensure the engagement team has the competencies required to operationalise the strategy”
7. “establish an action plan for meeting the desired policy objective”
8. “develop a monitoring and learning system” (Young & Mendizabal, 2009).

The Outcome Mapping Learning Community breaks their steps down into three stages, 12 steps:

INTENTIONAL DESIGN

1. “Vision”
2. “Mission”
3. “Boundary Partners”
4. “Outcome Challenges”
5. “Progress Markers”
6. “Strategy Maps”
7. “Organizational Practices”

OUTCOME & PERFORMANCE MONITORING

8. “Monitoring Priorities”
9. “Outcome Journals”
10. “Strategy Journal”
11. “Performance Journal”

EVALUATION PLANNING

12. “Evaluation Plan” (Hearn, Schaeffer & Ongevalle 2009)

Jones and Hearn (2009) write that outcome mapping works best when in partnerships, capacity building, in need of deeper social understanding, policy change, addressing complex issues, and in order to embed dialogue. The framework requires good timing, collaboration, flexibility, as well as a shift in perspective and mindset.

Mixed-Method / Alternative / Innovative Analysis

Professional evaluators have moved away from discipline-based evaluations employing a ‘‘single best method’’ and now tend to embrace mixed-method evaluation approaches in which decisions on methods follow efforts to understand the needs of information end-users (Horton and Mackay 2003). Horton (1986) argues that innovative forms of institutional analysis are necessary. He finds methods such as key informant interviewing and case studies to be needed for assessing changes in complex organisations but outside the bounds of traditional economics. Hence, he calls for professional contributions to be solicited from other disciplines, like anthropology, sociology and management.

Guijt (2007) cites several evolving ‘schools of thought’ (methods) for evaluating social change and innovation and explains that “creating an appropriate assessment and learning process requires mixing and adapting a combination of frameworks, concepts and methods to ensure they address the information and reflection needs and match existing capacities” (p 17). What each of the proposed methodologies have in common is a focus on structural issues that promote or hinder critical aspects of social innovation – gender equity, ethnic conflict, political representation, access to resources, etc.

Mama Cash, a Dutch organisation promoting social change envisioned by and directed toward women, established its own methodology for assessment. Called “Making the Case”, the methodology seeks to determine the extent to which: 1) the community views the issue(s) differently because of this project; 2) the community behaves differently because of this project; 3) individuals in the community engage more because of this project,; 4) policy has changed because of this project; and 5) the project has contributed to progress in spite of resistance (ibid.).

“*Accountability, Learning, Planning System*” is ActionAid’s methodology for assessment. The method is based on conceptual recognition that assessment, learning quality, and outputs are dependent on principles, attitudes, and behaviours. “Who wants to know what – and why it matters” is central to established procedural standards and directives. All organisational levels are requested to use stories, participatory review and reflection, reports, and external peer reviews to address: 1) what they did; 2) what rights were pursued; 3) what changes resulted for who—(a) critical thought or action, (b) tangible change in conditions, (c) growth in organisation or movement, (d) policy and practice—and; 4) what power relations were imparted (ibid.).

The *Society for the Promotion of Area Resource Centres* in India bases assessment on the central concept that the poor require space physically, emotionally, and socially in which to organize and develop. Therefore, the use of various assessment and learning methods (group exchange, stakeholder feedback, action learning/research, self evaluation, commissioned external perspectives, and individual professional development) are embedded into the action itself. The strength in this approach is common ownership of the process and idea of social change (ibid.).

Centro de Tecnologias Alternativas – Zona da Mata in Brazil embeds methodologies such as planning, diagnosis, monitoring, evaluation, and systematization—all participatory—into its work programmes. An internal and evolving learning process has involved all organisational levels since 2002 and has developed the following mix of approaches (Table 6; ibid.).

Table 6: Guijt (2007) Mechanisms for learning

What	*Frequency*	*Who participates*	*Focus*	*Outputs*
Regular contact between team members and other key actors, esp. farmer leaders	Daily	Farmers and team members	Whatever the issue is at hand	Quality of relationships, updated
Team meetings	Every 2 weeks	Team	Operational snags, decisions, short term planning	Decision, clarity about action points and responsibilities, support with problems
Annual reviews	Annual	Assembly with the team	Actual activities compared to planned	Shared evaluation of the year's achievements and problems
External monitor	Ongoing, new person every 3 years, annual commentary	External person and CTA team	If CTA is maintaining its planned focus and is learning	Reviews quality of organisational learning process to optimise and (re)align with organisational goals
Action research projects	Incidental	Varies	Topics have included: learning/M&E system; agroforestry; participatory municipal development	Insights about the questions being asked
Exchanges with other NGOs	Often but not regular	Team members and often farmers	Depends on topic selected (on technical, management issues or values)	New insights on operational issues, strategies, building of leadership
Systematising of lessons	Periodic, at the end of a programme or a phase	Farmers, CTA staff	Strategy, vision, leadership development	Insights that feed next phase of work or for wider sharing in the region/among NGOs
External evaluations	Every three years	External team with CTA	All operations and impacts	Key strategic challenges for current work
Courses for external groups or staff	One off, as needs emerge	Team members sometimes with specific farmer(s) leaders	Depends on need (e.g. gender, leadership or agroforestry training)	Shared conceptual and practical clarity, skills built, facilitation capacity enhanced

Source: Guijt, 2007

An evaluation of the impact of social networks carried out by IFPRI used a combination of SNA and economic estimation techniques under a single analytical framework. The authors argued that SNA offers tools for identifying important successes and key challenges to collective action. However, in order to promote adoption of innovative technologies, the study found that SNA must go beyond mapping the structure of such networks to identify local characteristics that may influence targeting, etc. (Matuschke 2008).

Other evaluations of agricultural research and extension services have focused on assessing economic returns on investment. Birkhaeuser et al. (1991) looked at 48 studies in 17 countries to assess the effect of distinct aspects of extension, including knowledge diffusion, adoption of improved technology and changes in agricultural productivity.

In regard to alternative monitoring and evaluation strategies, Guijt (2007) lists key concerns to consider. She says these strategies should exercise caution in:

- Identifying drivers of social change and underlying assumptions;
- Attributing beneficial impacts to specific interventions;
- Constructing and utilizing indicators or questions;

- Promoting skills for facilitating critical reflection and people's ability for assessment design and implementation;
- Assessing unequal power relationships as they relate to relationships, ethics, and standards; and
- Making generalizations based on assessment findings (ibid.).

Other Approaches to Impact Evaluation

There is a wide range of literature focused on evaluating and measuring change/impact upon beneficiaries. Some of this is relevant to the impact evaluations RIU is interested in. Of clear relevance, Krznaric (2007) includes a "rough guide to how change happens" (see Table 7) that could assist in framing evaluations generally or a theory of change in particular (pp 31-21).

Table 7: Krznaric (2007) Rough guide to how change happens

ASK YOURSELF…	REMEMBER TO THINK ABOUT…
1. What is the change that you want to explain? What was the situation before the change? And what was the situation after it?	Before-and-after contrasts in: • Behaviour of, and relations between, individuals and/or groups • Policies and practices of institutions (state, private sector, civil-society organisations) • Social attitudes and beliefs • The state of the natural environment • The state of human well-being
2. Who or what was involved in the change? Identify all relevant actors who: • were changed • were active agents of change • facilitated the change • resisted the change	People, as: • Individuals • Social groups (such as women workers, indigenous people) Institutions of: • State (executive, military, judiciary, bureaucracy, political parties) • Society (development NGOs, religious organisations, unions, the media) • Economy (corporations, small and medium-sized enterprises, informal workers, unpaid care-givers) • Global governance (WTO, World Bank, UN)
3. What strategies were used to bring about the change? Thinking about all the actors identified above: For the active agents of change, what strategies did they use? (maybe many at one time) For those who facilitated or supported change, did they also use strategies or take part in strategies? For those resisting change, what were their strategies of resistance? Why did they fail? Or did they partially succeed?	Strategies concerning individuals or (non-organised) groups: • Changing individual behaviour (using incentives or threats, making appeals to self-interest, altruism, fear, morality) • Reshaping worldviews/paradigms of understanding (through education, demonstration, dialogue, experience) • Promoting new human relationships through mutual understanding/empathy/reconciliation • Encouraging grassroots participation • Relying on charismatic and visionary leaders, highly networked individuals Strategies concerning institutions and organisations: • Pursuing reform or revolution? • Using coercion or consent? • Working inside or outside of the system? • Taking global/national/local action? • Working for short-term or long-term goals? • Taking legal or illegal action? • Following a top-down or bottom-up strategy? • Creating alliances or encouraging divisions? • Implementing blueprint plans or allowing diversity and experiment? • Forming organisations and movements • Using mass mobilisation/direct action/confronting/challenging • Focusing on organisational learning and flexibility • Empowering people • Providing assistance • Improving communication and information flow • Developing new technologies • Altering the social environment
4. What were the contexts that affected how the change happened? Which contexts *promoted* change?	State context: Regime type, military power, bureaucratic accountability, judicial autonomy, decentralisation, party fragmentation, corruption, history of state formation, political rights, civil war.

ASK YOURSELF…	REMEMBER TO THINK ABOUT…
Which contexts *permitted* change? Which contexts were *barriers* to change? What would a historian say? What would an economist say? A sociologist? Think with a different hat on… Which type of context (state, social, economic, global, environmental, or systemic) do you think was of little relevance to change? Now imagine that it *was* relevant, even important. How could that be so?	Social context: Worldviews and ideologies, nationalism, class structure, gender roles, family structure, religion, urbanisation and housing, cultural autonomy, civil liberties and media freedom, associational life, education/health access and levels, social stability, violence and crime, migration, colonial legacies, trust, social memory, social denial. Economic context: Industrialisation, property distribution, wealth inequality, privatisation, regulation, market access and distortions, corporate power and practice, labour conditions and laws, supply chains, access to technology, infrastructure, inflation, unemployment, macro-economic stability. Global context: Biased international trade rules, commodity price fluctuations, imperialist ambitions, balance of power, interstate conflict, terrorist threats, arms trade, effectiveness of UN institutions, efficacy of international law, Internet access. Environmental context: Geographic differences, resource availability/distribution, climate change, biodiversity, natural disasters, demographics, disease. Systemic context: Power relationships, interdependence, competition, inequality, historical precedents, uncertainty, chance, unknown factors.
5. What was the process or pathway of change? If you could 'draw' the change what would it look like? What kind of pathway did it follow? Change takes place on many levels, so there were probably many simultaneous 'pathways'. How did they interact?	Types of processes: • Cumulative progress/modernisation • Specific sequences, varying pathways, or cycles? • Tipping points • Demonstration effects • Crises or key junctures? • Conflicts or interactions? • Catalysts or cataclysms? • Internal or external pressures? • Systemic self-regulation • Unknown processes
6. Assessing the elements of change Of all the elements you have identified above, which would you pick out as the main ones that led to the change? Were any sufficient alone? Were they all necessary? How did they interact with each other? Which of them do you think is least acknowledged? And for all that has changed, what has not changed?	Some final considerations: • Most processes of change are extremely complex and defy single explanations • You may not have enough information to analyse the change effectively • Beware your personal assumptions and prejudices that affect your analysis • Take into account that your disciplinary speciality or training may still lead you to favour some explanations over others
Note: This is not a model of change; it is a rough guide to thinking through how change has happened in different contexts, from a range of disciplinary perspectives. It is a stimulus to thought rather than a blueprint to follow, a list of possible ingredients rather than a recipe. It makes no attempt to prioritise some aspects of change as being more important than others, nor does it specify how they may interact with each other, or over what time period they operate. The items in parentheses ( ) are examples, not an exhaustive list of the contents of a theme.	

Source: Krznaric, 2007

In thinking about the approach to impact evaluation design, the Foundation of Social Return on Investment (SROI) (2004) advises the following steps in determining the SROI:

- Setting Objectives and Identifying Outputs to Measure
- Decide Which Outcomes Can Be Monetised
- Find Data Necessary to Monetise Outcomes
- Calculate Value of Outcomes, Subtracting "What Would've Happened Anyway" (Impact)
- Calculate SROI based on Impact Divided by Investment (p 7.1)

Pertaining to identifying potential indicator elements, Parsons, Eoyang, Sherman and Williams (2007) attend to the Exploratory evaluation designed by University of New Hampshire's Food and Society initiative. This evaluation evaluated outcomes the initiative affected in policy, scholarship, farms/acreage, markets, and partners.

Regardless of the final evaluation framework, an important overall lesson from these approaches is stakeholder participation. In their 'Managing for Impact in Rural Development: A Guide for Project M&E' (2002), Guijt and Woodhill discuss how policy change itself can be a significant motivation for the primary stakeholders. Stakeholders should be made aware that the project they are involved with is aimed at influencing government policy and will enable "their voices to be heard at policy levels" (p 7.24).

CONCLUSION

It goes without saying that RIU's overall evaluation approach will be best informed by the considerations, reconsiderations, and recommendations found in this body of work. From this review of literature on the evaluation of systems and programmes, the following points would provide an especially useful guide as RIU continues to formulate impact evaluations.

In particular, EIARD (2006) has offered clearly stated standards of "good practice" with which to begin our evaluations: a clearly identified and described target object and context; a logical strategy superimposed upon a clearly stated model; a clear statement of objectives, scope, and limitations; a logical, transparent rationale for the plan completed before and continuing throughout evaluation; a statement of hypothesis as to expected impacts; a consideration of what other factors affect impacts; the assenting or dissenting rationale of other informed stakeholders.

Guijt (2007) makes suggestions to specifically to intermediaries involved with assessment between funders and social change organizations, such as the role RIU will placy. She recommends:

(1) integrating the evaluation as a part of rather than separate from the social change intervention

(2) taking a participatory action instead of implementation approach to the social change

(3) demonstrating the patience and time that open communication requires, especially where equality is a novel idea

(4) facilitating a capacity within the local group to identify assessment needs rather than directly expressing those needs

(5) clearly communicating concepts that effective remote assessment work requires

(6) adopting a questioning instead of a directing position

(7) while bringing an objective—outsider's—perspective, giving equal attention to local assessment perspectives

(8) accepting that some quality, depth, and time must be compromised in strict funding contexts

(9) developing opportunities for others to safely reflect and honestly assess, even if this requires initial self-deprecation in order to demonstrate an openness to critique

(10) maintaining consistent connections with those directly involved in change.

Summarizing other papers in the same issue of publication, Horton and Mackay (2003) identify several important considerations for enhancing the overall benefit of evaluations. These include:

- Assessments and evaluations should be kept separate from research, considering the former to be complete only when utilized by decision-makers.
- Stakeholders should be included in the process of assessment and evaluation.
- In addition to measuring variables, impact evaluations should address how and why organisations and institutions have changed.
- Evaluations should address the broader complexity of the agricultural innovation system from a pro-poor perspective.
- Evaluations should model the pathway between programme logic and impact that contributes to the design of research, evaluations, and understanding how efforts in research and development are created.
- Evaluations should include achievements that are important to stakeholders instead of focusing only on easily measured achievements.
- Evaluations should draw evidence from the full range of stakeholder groups that results in findings that transcend a limited point of view.
- Frame evaluations using the constructs of many disciplines to address the broader, more complex processes that are of concern to stakeholders.
- Equip stakeholders to address their responsibilities by identifying and seeking ways for them to use results early on.
- Select methods that are appropriate for the circumstances stakeholders represent rather than an overall best method.
- Develop the multidisciplinary competency of staff and the means for organisations to productively utilize this competency. And have realistic expectations about causal complexity.

DOCUMENTS REVIEWED

Adato, M. and R. Meinzen-Dick (2002) 'Assessing the Impact of Agricultural Research on Poverty Using the Sustainable Livelihoods Framework', International Food Policy Research Institute (IFPRI) Food Consumption and Nutrition Division (FCND) Discussion Paper 128, Environment and Production Technology Division (EPTD) Discussion Paper 89, Washington DC (http://www.ifpri.org/sites/default/files/pubs/divs/eptd/dp/papers/eptdp89.pdf).

Agriculture and Rural Development (ARD) (2006) 'Enhancing Agricultural Innovation: How to Go Beyond the Strengthening of Research Systems', The International Bank for Reconstruction and Development, Washington DC (http://siteresources.worldbank.org/INTARD/Resources/Enhancing_Ag_Innovation.pdf).

Active Learning Network for Accountability and Performance (ALNAP) and Channel Research (2010) 'Evaluation of Humanitarian Action (EHA): Course Reference Manual' (Version 1.30), London UK (www.alnap.org/pool/files/2010-training.pdf).

Anderson, J. R. (2007) 'Agricultural Advisory Services', Background Paper for the World Development Report 2008, Agriculture and Rural Development Department, World Bank, Washington DC (http://siteresources.worldbank.org/INTWDR2008/Resources/2795087-1191427986785/Anderson_AdvisoryServices.pdf).

Birkhaeuser, D., R. E Evenson and G. Feder (1991) 'The Economic Impact of Agricultural Extension: A Review', *Economic Development and Cultural Change*, Vol 39 (No 3) pp 607-50.

Birner, R., K. Davis, J. Pender, E. Nkonya, P. Anandajayasekeram, J. Ekboir, A. Mbabu, D. Spielman, D. Horna, S. Benin and M. Cohen (2006) 'From "Best Practice" to "Best Fit": A Framework for Analyzing Pluralistic Agricultural Advisory Services Worldwide', Development Strategy and Governance Division Discussion Paper No. 37, IFPRI, Washington D.C. (http://www.ifpri.org/DIVS/DSGD/dp/dsgdp37.asp).

Center for Global Development (CGD) (2006) 'When Will We Ever Learn? Improving Lives through Impact Evaluation'. Report of the Evaluation Gap Working Group, Washington DC, (www.cgdev.org/files/7973_file_WillWeEverLearn.pdf).

Chambers, R., D. Karlan, M. Ravallion and P. Rogers (2009) 'Designing impact evaluations: different perspectives', The International Initiative for Impact Evaluation (3ie) Working Paper 4, Washington DC (http://www.3ieimpact.org/admin/pdfs_papers/50.pdf).

Conley, T. and C. Udry (2001) 'Social learning through networks: The adoption of new agricultural technologies in Ghana', *American Journal of Agricultural Economics*, Vol 83 (No 3) pp 668–673.

Connell, J. P. and A. C. Kubisch (1998) 'Applying a Theory of Change Approach to the Evaluation of Comprehensive Community Initiatives: Progress, Prospects, and Problems', In K. Fulbright-Anderson, A. C. Kubisch, J. P. Connell (Eds.) *New Approaches to Evaluating Community Initiatives, Vol. 2: Theory, Measurement, and Analysis*, The Aspen Institute, Washington DC, pp 15-44 (https://communities.usaidallnet.gov/fa/system/files/Applying+Theory+of+Change+Approach.pdf).

Department for International Development (DfID) (2008) 'Research Strategy 2008-2013', London UK (http://webarchive.nationalarchives.gov.uk/+/http://www.dfid.gov.uk/pubs/files/Research-Strategy-08.pdf).

Ekboir, J. M., G. Dutrénit, G. Martínez V., A. T. Vargas and A. O. Vera-Cruz (2009) 'Successful Organizational Learning in the Management of Agricultural Research and Innovation: The Mexican Produce Foundations', IFPRI Report 162, Washington DC (http://www.ifpri.org/sites/default/files/publications/rr162.pdf).

European Initiative for Agricultural Research for Development (EIARD) (2006) 'Impact Assessment and Evaluation in Agricultural Research for Development', The Task Force on Impact Assessment and Evaluation, London UK (http://www.eiard.org/docu/Eiard_Ia_paper.html).

Foundation of Social Return on Investment (2004) 'Measuring social impact', London UK (http://sroi.london.edu/Measuring-Social-Impact.pdf).

Guijt, I. (2007) 'Assessing and Learning for Social Change: A Discussion Paper', Institute of Development Studies (IDS), Brighton UK (http://www.ids.ac.uk/download.cfm?objectid=3DF43790-5056-8171-7BE05D04AD0A1C8C).

Guijt, I. and J. Woodhill (2002) 'Managing for Impact in Rural Development: A Guide for Project M&E', International Fund for Agricultural Development, Rome IT (http://www.ifad.org/evaluation/guide/index.htm).

Hall, A., L. Mytelka and B. Oyeyinka (2006) 'Concepts and guidelines for diagnostic assessments of agricultural innovation capacity', United Nations University - Maastricht Economic and social Research and training centre on Innovation and Technology Working Paper 017, Maastricht NL (http://www.merit.unu.edu/publications/wppdf/2006/wp2006-017.pdf).

Hearn, S., H. Schaeffer and J. V. Ongevalle (Eds) (2009) 'Making Outcome Mapping Work - Volume 2: Innovations in Participatory Planning, Monitoring and Evaluation', Outcome Mapping Learning Community, London UK (http://www.outcomemapping.ca/download.php?file=/resource/files/simonhearn_en_Making%20Outcome%20Mapping%20Work_online.pdf).

Horton, D. and R. Mackay (2003) 'Using Evaluation to Enhance Institutional Learning and Change: Recent Experiences with Agricultural Research and Development', *Agricultural Systems*, Vol 78 (No 2) pp 127-142.

Horton, D. (1986) 'Assessing the Impact of International Agricultural Research and Development Programs', *World Development*, Vol 14 (No 4) pp 453-468.

Horton, D. (1998) 'Disciplinary roots and branches of evaluation: some lessons from agricultural research', *Knowledge and Policy*, Vol 10 (No 4)pp 31–66 (ftp://ftp.cgiar.org/isnar/publicat/discuss/dp96-7.pdf).

Hovland, I. (2007) 'Making a difference: M&E of policy research', Overseas Development Institute (ODI) Working Paper 281, London UK (http://www.odi.org.uk/resources/download/1751.pdf).

Independent Evaluation Group of the World Bank (IEG) (2006) 'Impact Evaluation: An overview and some issues for discussion', Organization for Economic Cooperation and Development

(OECD) (DAC) Room Document 5, London UK (http://www.cgdev.org/doc/eval%20gap/HowardWhiteDAC.pdf).

Jones, H. and S. Hearn (2009) 'Outcome Mapping: a realistic alternative for planning, monitoring and evaluation', ODI Background Note, London UK (http://www.odi.org.uk/resources/download/4118.pdf).

Jones, N., H. Jones, L. Steer and A. Datta (2009) 'Improving impact evaluation production and use', ODI Working Paper 300, London UK (http://www.odi.org.uk/resources/download/3177.pdf).

Krishnan, P. and E. Sciubba (2004) 'Links and Architecture in Village Networks' (previously 'Endogenous Network Formation and Informal Institutions in Village Economies'), Cambridge Working Paper in Economics 0462, University of Cambridge, Cambridge UK (http://www.econ.cam.ac.uk/dae/repec/cam/pdf/cwpe0462.pdf).

Kristjanson P., P. K. Thornton, B. Douthwaite, J. Watts, R.S. Reid and L. Withers (2002) 'Evaluating Impact of Agricultural Research And Development - Future Harvest Research Centre Approaches', presented at the African Evaluation Association Conference, Nairobi KE (http://www.researchintouse.com/learning/learning40innovnarra.html)

Kusek, J.Z. and R.C. Rist (2004) 'Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners', The World Bank, Washington DC (http://www.oecd.org/dataoecd/23/27/35281194.pdf).

Krznaric, R. (2007) 'How Change Happens: Interdisciplinary Perspectives for Human Development', Oxfam GB Research Report, Boston MA (http://www.oxfam.org.uk/resources/issues/education/downloads/research_change.pdf).

Lele, U., J. Pretty, E. Terry and E. Trigo (2010) 'Transforming Agricultural Research for Development' (Draft 6.0), Report for the Global Conferences on Agricultural Research for Development, Montpellier FR (http://www.ciat.cgiar.org/cgiar/Documents/transforming_agricultural_research_11mar10.pdf).

Matuschke, I. (2008) 'Evaluating the Impact of Social Networks in Rural Innovation Systems: An Overview', IFPRI discussion paper, Washington DC (http://www.ifpri.org/sites/default/files/publications/ifpridp00816.pdf).

Munshi, K. (2004) 'Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution', *Journal of Development Economics*, Vol 73 (No 1) pp 185–213.

Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC) (2002) 'Glossary of Key Terms in Evaluation and Results-based Management Proposed Harmonized Terminology', OECD/DAC, Paris FR (www.oecd.org/dataoecd/14/31/17484948.pdf).

Office of Juvenile Justice and Delinquency Prevention (2010) 'What is a CCI?', Washington DC, (http://www.ccitoolsforfeds.org/getting_what.asp)

Pant, L. P. and H. H. Odame (2008) 'Assessing Social Innovations in Agricultural Research and Development Partnerships', Delivered at Workshop on Rethinking Impact: Understanding the Complexity of Poverty and Change, Cali CO

(http://www.prgaprogram.org/riw/files/papers/Assessing%20Social%20Innovations_Pant%20and%20Hambly%20Odame.pdf).

Parsons, B., G. Eoyang, J. Sherman and B. Williams (2007) 'Designing Initiative Evaluation: A Systems-oriented Framework for Evaluating Social Change Efforts', T. Behrens, H. Yang and L. Helstowski (Eds) W. K. Kellogg Foundation, Battle Creek MI (http://gametlibrary.worldbank.org/FILES/995_%20Systems%20thinking%20in%20evaluation%20-%20Kelloggs.pdf).

Patel, S. (2007) 'Reflections on Innovation, Assessment and Social Change Processes: A SPARC Case Study, India', Case study produced for the 'Assessing Social Change' initiative, IDS, Brighton UK (http://www.ids.ac.uk/ids/Part/proj/socialchange.html).

Patton, M. Q. and D. Horton (2009) 'Utilization Focused Evaluation for Agricultural Innovation', Institutional Learning and Change Brief 22, Maccarese IT (http://ageconsearch.umn.edu/bitstream/52533/2/ILAC_Brief22_Utilization_Focus_Evaluation.pdf).

Perrin, B. (2000) 'How to — and How Not to — Evaluate Innovation' (Revised 14 May 2001), Presentation to the UK Evaluation Society Conference, London UK (http://www.mande.co.uk/docs/perrin.htm).

Pestieau, C. (2003) 'Evaluating Policy Research', Research Paper W/22, Canadian Policy Research Networks Inc., Ottawa CA (http://www.bibliotheque.assnat.qc.ca/01/mono/2004/07/778722.pdf).

Peterson, W., G. Gijsbers and M. Wilks (2003) 'An Organizational Performance Assessment System for Agricultural Research Organizations: Concepts, Methods, and Procedures', The Hague: International Service for National Agricultural Research, Washington DC (ftp://ftp.cgiar.org/isnar/Publicat/rmg7/rmg7.pdf).

Research Into Use (RIU) (2010) Website, London UK (http://www.researchintouse.com/learning/learning40innovnarra.html)

Romani, M. (2003) 'The impact of extension services in times of crisis: Cote d'Ivoire (1997-2000)', CSAE WPS/2003-07, Centre for the Study of African Economies, University of Oxford, Oxford UK (http://129.3.20.41/eps/dev/papers/0409/0409053.pdf).

Start, D. and I. Hovland (2004) 'Tools for Policy Impact: A Handbook for Researchers' (Reprinted 2007), ODI, London UK (http://www.odi.org.uk/resources/download/156.pdf).

Sulaiman V, R. (2010) 'Changing Contours of Research And Development: Lessons From South Asia', Research Into Use, London UK (http://www.researchintouse.com/resources/riu10lessons-asia-sulaimanv.pdf).

Waddington, H., B. Snilsveit, H. White and J. Anderson (2010) 'The Impact of Agricultural Extension Services', 3ie Synthetic Reviews – SR009, 3ie, Washington DC (http://www.3ieimpact.org/admin/pdfs_synthetic/009%20Protocol.pdf).

White, H. (2009a) 'Some Reflections on Current Debates in Impact Evaluation', 3ie Working Paper 1, Washington DC (http://www.3ieimpact.org/admin/pdfs_papers/11.pdf).

White, H. (2009b) 'Theory-Based Impact Evaluation: Principles and Practice', 3ie Working Paper 3, Washington DC (http://www.3ieimpact.org/admin/pdfs_papers/51.pdf).

Young, J. and E. Mendizabal (2009) 'Helping researchers become policy entrepreneurs: How to develop engagement strategies for evidence-based policy-making', ODI Briefing Paper 53, London UK (http://www.odi.org.uk/resources/download/1127.pdf